UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

DTIC ELECTE D

AD-A215 084

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited
	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 89-1489
6a. NAME OF PERFORMING ORGANIZATION: Department of Psychology	6b. OFFICE SYMBOL (If applicable)
7a. NAME OF MONITORING ORGANIZATION: Air Force Office of Scientific Research/NL	
6c. ADDRESS (City, State, and ZIP Code): University of Arizona, Tucson, Arizona 85721	7b. ADDRESS (City, State, and ZIP Code): Building 410, Bolling AFB, D.C., 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR	8b. OFFICE SYMBOL (If applicable): NL
9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR-89-0075	
8c. ADDRESS (City, State, and ZIP Code): Building 410, Bolling AFB, D.C., 20332-6448	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO.
61102F	2313	A4	

11. TITLE (Include Security Classification)
Measures of Subjective Variables in Visual Cognition

12. PERSONAL AUTHOR(S) Mary A. Peterson

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Annual Technical	FROM 9/1/88 TO 9/30/89	9/30/89	95

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
05	10	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Visual cogniton; visual perception; attention; elective set; figure-ground organization; reversible figures; imagery ; (KT)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A series of 10 experiments examined the subjective variables involved in the perceptual organization of shapes and objects; in particular, the role played by (1) perceptual intentions, (2) structural knowledge, and (3) spatial attention. A series of experiments identified functional consequences of structural knowledge in both perception and imagery and functional differences between the two types of of perceptual selectivity we examined. We found that (1) prototypical shapes are perceived faster than quasitypical shapes (the goodness of the set between structural representations and these two types of shapes was assumed to differ), (2) structural knowledge regarding shape components is a necessary requirement for imagery reversal, (3) perceptual intentions are more effective when directed to prototypical rather than nonprototypical shapes, implicating structural knowledge as a mechanism through which perceptual intentions operate, and (4) spatial attention can be directed to parts of objects whereas perceptual intentions may operate holistically. In addition, the experiments provided some evidence that multiple shape representations are activated prior to shape recognition, and led to a new model of figure-ground organization. Keywords

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL: Alfred R. Fregly, Ph.D.	22b. TELEPHONE (Include Area Code): (202) 767-5021 — 22c. OFFICE SYMBOL: NL

DD Form 1473, JUN 86 — Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

89 11 20 050

Annual Report

<u>Measures of Subjective Variables in Visual Cognition</u>

9/1/88 - 9/30/89

Mary A. Peterson

University of Arizona

Table of Contents

Annual Report
Measures of Subjective Variables in Visual Cognition
9/1/88 - 9/30/89

Mary A. Peterson
University of Arizona

The research conducted in my laboratory during the first year of joint NSF/AFOSR support examined the subjective variables involved in the perceptual organization of shapes and objects; in particular, the role played by (1) perceptual intentions, (2) structural knowledge, and (3) spatial attention. A series of experiments identified functional consequences of structural knowledge in both perception and imagery and functional differences between the two types of perceptual selectivity we examined. We found that (1) prototypical shapes are perceived faster than nonprototypical shapes (structural knowledge about these two types of shapes was assumed to differ, see below), (2) structural knowledge regarding shape components is a necessary requirement for imagery reversal, (3) perceptual intentions are more effective when directed to prototypical rather than nonprototypical shapes, implicating structural knowledge as a mechanism through which perceptual intentions operate, and (4) spatial attention can be directed to parts of objects whereas perceptual intentions may operate holistically. In addition, the experiments provided some evidence that multiple shape representations are activated prior to shape recognition, and led to a model of figure-ground organization, which raises many important questions about perceptual organization.

The experiments that have been conducted fall into two general categories: perception experiments and imagery experiments. In what follows, I will summarize the rationale, methods, results, and implications for the perception experiments first and then for the imagery experiments.

I. Perception experiments: Intention and structural knowledge

In previous research Hochberg and I have examined a type of selectivity that has been relatively neglected in the field of visual cognition: the viewer's perceptual intentions to match one potential interpretation to a stimulus. Using a procedure called the "opposed-set procedure" in which we instructed observers to try to hold one interpretation of a reversible figure (Peterson & Hochberg, 1983), we have shown that the viewers' perceptual intentions can influence the perceived organization of partially biased Necker cubes. The reports made by subjects following opposed-set instructions are not simply responses to the demand

character of the instructions; instead, they reflect what the viewers perceive (Hochberg & Peterson, 1987; Peterson, 1986). Other evidence suggests that the intention effects are not mediated by eye movements: neither saccadic movements nor convergence changes covary with instructions in the opposed-set procedure (Peterson, 1984; 1986), thereby leaving open a role for more central cognitive mechanisms, which remain to be explicated.

One possible mechanism through which the viewers' intentions might be instantiated is through priming of the mental representation corresponding to the intended organization. Priming would act to increase the likelihood that the intended representation will be matched to the stimulus. We examined a specific version of a priming hypothesis -- that top-down activation operates on a representation that entails knowledge about the shape parts and their spatial relationships (i.e., structural knowledge). There are a number of reasons to expect that instructed intentions operate by priming structural knowledge as opposed to either semantic knowledge or knowledge about individual features. First, priming studies have shown that priming of pictures from words is quite weak (Roediger & Weldon, 1987), perhaps because the prior presentation of a word primes semantic knowledge rather than the structural knowledge necessary for picture identification. Second, in most current theories of perception, knowledge of the parts or features comprising a shape is not sufficient for identification because many shapes are constructed from the same component parts. Knowledge of the spatial relationships among parts is necessary to differentiate among a number of alternatives (Biederman, 1987; Marr, 1982; Marr & Nishihara, 1978).

A number of predictions follow from the hypothesis that structural knowledge is mediating the intention effects. One prediction is that the intention effects should increase as structural knowledge increases. To examine this question, we manipulated structural knowledge by using <u>prototypical</u> shapes and <u>nonprototypical</u> shapes. <u>Prototypical</u> shapes are shapes that unambiguously represent the central tendency of a class of shapes (i.e., most observers asked to describe the shape label it with the same unelaborated label); <u>nonprototypical</u> shapes are shapes that ambiguously represent a specific member of a class of shapes (i.e, observers asked to describe the shape tend not to agree on any single label, and the labels they do assign tend to be elaborated by a number of descriptors). We assume that observers have more structural knowledge about prototypical shapes, because representations of non-prototypical shapes may entail semantic elaboration of the structural knowledge or may be constructed from unstable parts. (We do not use Rosch's terminology of basic,

subordinate, and superordinate levels, because those distinctions may not describe perceptual organization per se, but only naming. Nor do we distinguish between "good" shapes and "poor" shapes, because those labels are anchored in stimulus dimension differences (cf. Palmer, 1975). By using figure-ground stimuli, we were able to hold stimulus dimensions constant while varying the prototypicality of the shape. The figure-ground experiments were conducted in collaboration with Erin Harvey, Brad Gibson, and Hollis Weidenbacher.

Prototypicality and figure-ground organization.

We chose to use a figure-ground stimulus because the contour in figure-ground stimuli defines two distinct shapes, only one of which can be seen at a time. Consider the classic Rubin figure shown in Figure 1. When the contour is assigned to the white region, the faces are seen and the black region appears contourless, becoming the ground against which the faces are viewed. When figure and ground reverse so that the black region appears to be figure, the vase is seen, but the faces are not. The white region is now contourless and appears to continue behind the vase. Because figure-ground stimuli are reversible and because the two shapes competing for the contour can be chosen to vary in prototypicality, figure-ground stimuli presented ideal stimuli for asking whether perceptual intentions directed to a prototypical shape were more effective than perceptual intentions directed to a nonprototypical shape.

The figure-ground stimuli we used are shown in Figures 2 and 3. The stimuli were biased toward the center-as-figure interpretation by the variables of area, closure, and symmetry (Koffka, 1935), and 88% of observers shown these stimuli reported seeing the center as figure. As can be seen in Figure 2, the center region is a more prototypical shape than the the surround. For example, observers reliably called the center of the figure on the left a vase or an architectural pediment. Although the center region cannot be said to be completely unambiguous, it was more prototypical than the shapes seen in the surround. When the shapes are rotated by 180 degrees, however, as they are in Figure 3, the surround is seen to be more prototypical than the center. The surround in the figure to the left represents two half silhouettes of women shown head to toe and the surround on the right represents two schematic profiles of men. The stimulus variables of area, symmetry, and closure do not change when these stimuli are inverted. Nor do the parts into which the contour would be parsed by some rule like locating the minima of curvature (cf. Hoffman & Richards, 1984). The only change with

inversion is in subjective variables such as the prototypicality of the two regions.

Methodology.

In Experiments 1 - 3, these stimuli were used with opposed-set instructions. A fixation point was placed in the center of the black region. Observers were asked to fix their eyes on this point for the duration of each 30 second trial. On alternate trials, they were instructed to try to hold the center or the surround as figure. They pressed one of two keys to indicate which region appeared to be figure at any moment during the trial, and the durations that each key was pressed were cumulated across the trial. The data I report below are the mean durations that an individual percept lasted. (The results are replicated with total durations.) I will report two measures: First, the mean duration that a given region was seen as figure will be evaluated as evidence whether, regardless of the viewers' intentions, prototypical shapes are seen as figure longer than non-prototypical shapes. Second, the difference in the mean duration that a region was seen as figure when it was the intended region (I) versus when it was the unintended region (U) will be our measure of the effects of intention. If I - U > 0, that will show that intention was effective. If I - U is greater for prototypical shapes than for nonprototypical shapes, that will show that intention is more effective on prototypical shapes, and will implicate structural knowledge as a mechanism through which intention operates.

Experiment 1

In the first part of Experiment 1, 16 subjects viewed the figure ground stimuli in the orientation shown in the top row of Figure 2 (hereafer called the inverted orientation), without having seen the upright stimuli and without any instructions regarding the shapes that might be found in the center or the surround. They participated in two opposed-set trials with each stimulus (one with each hold instruction). Between parts 1 and 2 of the experiment, the subjects were separated into two groups. The experimental group was shown upright versions of the stimuli, and the prototypical interpretations were pointed out to them in detail. This group viewed the upright stimuli in part 2 of the experiment. The practice control group did not view the upright stimuli at any time; during part 2 of the experiment, they continued to view the inverted stimuli. We predicted that if intention operated through structural knowledge, observers in the experimental group should be more successful at holding the surround as figure in part 2 than in part 1, and observers in the

practice control group should show little change in their ability to hold the surround as figure from part 1 to part 2.

Results. The results, shown in Table 1, were as predicted. The observers in the experimental group showed substantial increases in the mean duration that they succeeded in holding the surround as figure when it represented a prototypical shape, whereas the observers in the control group showed little change in their ability to hold the surround in part 2. In addition, the mean duration that observers in the experimental group were able to hold the center as figure decreased when the the surround represented a prototypical shape, a point to which we will return (see pp. 6-7). Intentions to hold the prototypical shapes were more effective than intentions to hold the nonprototypical shapes, as shown in Table 1, implicating structural knowledge as a mechansim through which intention operates.

Experiments 2A and 2B

Experiment 1 does not allow us to separate the effects of semantic knowledge from those of structural knowledge. We separated those variables in Experiments 2A and 2B. In Experiment 2A, we examined whether the increased ability to hold the surround as figure occurred because the observers were able to identify the individual parts of the shapes, or to assign a label to the surround, or whether the relative locations of the parts were critical as well. To that end, we used shapes constructed from the same parts as the upright shapes in Experiment 1, but in Experiment 2A, their relative locations were scrambled.

In part 1 of Experiment 2A, a group of naive 16 observers viewed the upright scrambled shapes without knowledge regarding the prototypical shape from which they were drawn. Between parts 1 and 2, observers were divided into two groups. The experimental group was shown the upright figures in their unscrambled form, and the prototypical interpretation for each was pointed out to them, as was the correspondence between the parts of the upright figures and the parts of the scrambled figures. (Each of the scrambled parts now had a name: e.g., "skirt" or "nose".) If structural knowledge regarding the shape's parts and their relative locations mediates the prototypicality effects and the prototypicality-dependent intention effects, and not simply knowledge about the identity of each part, the experimental group in Experiment 2A should not show the prototypicality effects shown by the experimental group in Experiment 1.

In Experiment 2B, we examined whether the upright orientation of the prototypical shape was critical to the prototypicality effects, or whether knowledge about the prototypical interpretation was sufficient. A group of 16 naive

observers viewed the inverted shapes in the first half of the experiment without any knowledge regarding the prototypical interpretation. Between parts 1 and 2, all observers were shown the upright stimuli and the prototypical interpretations were pointed out. In part 2, half of the observers viewed upright stimuli and the other half viewed inverted stimuli.

Experiment 2B provides another opportunity to examine whether structural knowledge per se mediates the intention effects, or whether they are mediated by semantic knowledge. Corballis (1988) has recently proposed that shapes can be identified regardless of their orientation by accessing a nonstructural representation, and that the time taken to access this representation does not vary with orientation. If a nonstructural representation with these properties is mediating the intention effects, the orientation of the prototypical shape should not matter.

Results. The results of Experiment 2A are shown in Table 2. Knowledge about the prototypical shape from which the scrambled parts were drawn (and hence, the identifiability of individual parts) did not increase the subjects' ability to hold the surround as figure: The observers in the experimental group were no more successful at holding the surround as figure than the observers in the control group, nor were their intentions to hold the surround as figure more effective than those of the observers in the control group.

The results of Experiment 2B are shown in Table 3. Knowledge about the prototypical shape that can be seen in the surround did not increase the mean duration that the surround was seen as figure when it was inverted. Nor were the intentions of observers in the experimental group any more effective than those of observers in the control group.

Discussion

The fact that the prototypicality effects were not obtained with scrambled stimuli or with inverted stimuli supports the hypothesis that perceptual intentions may operate through structural, rather than semantic, knowledge. The increased duration of surround-as-figure reports for the prototypical surround on "hold surround" trials shows that the likelihood of maintaining a region as figure is higher when that region represents a prototypical shape. In addition, the decreased duration of center-as-figure reports on the "hold center" trials in the experimental group of Experiment 1 shows that the likelihood of obtaining a prototypical region as figure by

reversal from the nonprototypical center-as-figure interpretation is also higher.

A theory of figure-ground reversal, intention, and prototypicality.

Preattentive opportunities for reversal. We account for the increased likelihoods of both obtaining and maintaining the prototypical region as figure by proposing the existence of intermittent preattentive opportunities for reversal. When an opportunity for a reversal occurs, a parallel search is initiated for shape matches to both sides of a contour. Whichever representation first reaches some criterion activation level is perceived. If the currently perceived shape is reconfirmed faster than the alternative, it is maintainted and no reversal occurs. If the alternative shape is confirmed faster than the currently perceived shape, a figure-ground reversal occurs.

Our proposal that a contour is parsed from both sides simultaneously prior to figure-ground reversal is new. Most theories assume that the figure is chosen by preattentive visual routines (cf. Hoffman & Richards, 1984; Ullman, 1984) before the contour is parsed into parts by some rule like locating the minima of curvature as identified from inside the figure (Hoffman & Richards, 1984), yielding a structural description that is unique to the shape on the side of the contour designated as figure. Recognition then follows a parallel search in which that structural description is mapped to multiple memory repesentations simultaneously. Our results, which suggest that the contour may be parsed from both sides simultaneously, imply a larger search through memory representations, and imply that the same contour can be used to construct two different structural descriptions simultaneously. The structural descriptions differ becasue the minima of curvature, and hence, the parts, identified from the left side of the contour differ from those identified from the right side.

Prototypicality and speeded matches. We explain the finding that the prototypical shapes are more likely to be both obtained and maintained by proposing that the mapping between the structural description and the memory representation is completed faster for the prototypical shapes than for the nonprototypical shapes (either because the structural knowledge about prototypical shapes is greater, or because the parts are represented in similar neighborhoods (cf. Riddoch & Humphreys, 1987)). Therefore, whenever an opportunity for a reversal occurs, and a simultaneous search is conducted for the shapes on both

sides of the contour, the prototypical shape will be more likely to win the contour.

Intentions, prototypicality, and orientation. We account for the finding that perceptual intentions were more effective when directed to the prototypical shapes than to the nonprototypical shapes by proposing that intention operates as the top-down activation of structural knowledge about the shapes, which further increases the likelihood of obtaining the prototypical shapes. Our finding that the prototypicality effects disappeared when the prototypical shape was inverted suggests that the structural description may not be matched through multiple orientation mappings simultaneously, as is assumed in many theories (Corballis, 1988; McClelland & Rumelhart, 1981), but may be mapped in a serial, or a partially serial, manner. The prototypicality-dependent intention effects disappeared as well, suggesting that teh location of the shape's top and bottom may be specified in the primed representation. These possibilities are explored further in Experiment 3, but before introducing that experiment, we discuss an alternative interpretation for the results of the preceding experiments.

An alternative account of the results of Experiments 1 - 2. There is an alternative explanation for the results of the previous experiments. Once observers were aware of the prototypical interpretation for the surround, they may have allowed their intentions or their spatial attention to stray to the surround so that they could see the more prototypical shape more often. Indeed, in some models of figure-ground organization, the focus of spatial attention determines which region will be seen as figure. This spatial attention explanation can account for the prototypicality effects obtained with the upright figures in Experiment 1, but it cannot explain the absence of prototypicality effects in Experiment 2B, unless one assumes that the observers in the experimental group of Experiment 2B simply continued using the strategy they had used in part 1 rather than using a strategy appropriate for the prototypical shape. Experiment 3 was conducted in order to distinguish between the parallel search account and the spatial attention account of the results of the previous experiments.

Experiment 3

To distinguish between the two interpretations, we sought a means to slow the speed of the match between the prototypical shape and its memory representation, but to leave unchanged its apparent prototypicality, and hence, its ability to attract attention. The manipulation we chose was to present the stimuli in five different orientations between upright and inverted, in

45 degree steps. Before viewing the stimuli, all subjects were shown the upright figures and the prototypical shapes were pointed out. If the results of Experiment 1 were due to lapses of intention or of spatial attention caused by knowing that the surround is more prototypical than the center, the prototypicality effects should be found at all orientations in this experiment, since observers were not given any practice seeing the surround as some shape other than the prototypical shape.

If the match between structural description and memory representation proceeds through different orientation maps in a serial (or partially serial) fashion (and there is some experimental support for this assumption (Jolicoeur, 1985)), then misorienting the prototypical shape should incrementally slow the match between the surround and the mental representation. If this manipulation results in a incremental decrease in the prototypicality effects as misorientation from upright increases, that would support both the simultaneous search hypothesis and the hypothesis that prototypical shapes are matched to memory representations faster than nonprototypical shapes. If the probability-dependent intention effects decrease with increasing misorientation from canonical upright, that would suggest that the structural knowledge primed by the intention instructions includes specification of a canonical coordinate axis.

Results and discussion. As shown in Figure 4, the results of Experiment 3 are consistent with the simultaneous search hypothesis. The prototypicality effects decreased incrementally as misorientation from upright increased. This supports the hypothesis that, when an opportunity for a reversal occurs, a speeded search is conducted for the two shapes constructed from the parts identified on both sides of the contour. As long as the prototypical shape is misoriented from the upright by less than about 60 degrees, the matching process for the prototypical shape is completed faster than the matching process for the nonprototypical shape. Misorienting the figure-ground stimulus slows the match to the representation of the prototypical shape sufficiently so that the prototypicality effects disappear somewhere around 45 - 60 degreees from upright. The intention effects follow a function very similar to the prototypicality effects; the prototypicality-dependent intention effects disappear at approximately the same point as the prototypicality effects.

The intention effects decreased as misorientation from upright increased as well, suggesting that intention may operate by priming a canonical representation of the intended shape.

Thus, top-down activation may not influence the matching process until the structural description is matched to a canonical memory representation. That is, intention does not appear to speed the orientation-normalizing process.

Remaining issues.

These experiments raise a number of experimental and theoretical questions. For example: What constitues an opportunity for a reversal? Are contours in the visual field always parsed from both sides simultaneously, or only when an opportunity for a reversal occurs? Does the introduction of relative depth into the display eliminate the simultaneous search for shapes on both sides of a contour? How do stimulus variables like area, symmetry, and closure interact with the prototypicality effects? Does prototypicality influence the first percept or only the likelihood of reversal? (An experiment directed to this question is currently in progress using brief masked exposures.) Is the presence of some minimum number of shape parts necessary for the occurrence of the prototypicality effects? Can the differences between prototypical shapes and nonprototypical shapes be explicated any better?

II. Spatial attention

Experiment 3 ruled out spatial attention as an explanation for the prototypicality effects. But that does not imply that spatial attention has no influence on perceptual organization, or that intention does not operate through spatial attention. The observers were successful at following instructions to try to hold the nonprototypical shapes as figure in Experiments 1 - 3 were successful. How is intention having an effect on nonprototypical shapes? One possiblity is that intention can also operate via the allocation of spatial attention. We examined the relationship between spatial attention and perceptual intention in Experiments 4 - 7.

Experiment 4

In this experiment, we examined whether allocating spatial attention to one region in a figure-ground stimulus increased the likelihood that that region would be seen as figure. In addition, we examined whether the effects of spatial attention vary with the prototypicality of the region to which attention was directed. We used figure-ground stimuli in which the area, symmetry, and closure of the regions on either side of the figure-ground contour were equated. The prototypicality of the shapes on either side of the contour were not equated, however: there was a prototypical shape on one side of the contour and a

nonprototypical shape on the other side. Observers were asked to fixate a fixation dot on either the left shape or the right shape and to try to hold either the left region or the right region as figure.

Results. We replicated the prototypicality effects obtained in Experiments 1 and 3: The mean duration that observers succeeded in holding the prototypical region was longer than the mean duration they succeeded in holding the nonprototypical region as figure, as shown in Table 5. In addition, we obtained a main effect of fixation/attention point: The mean durations that observers reported seeing a region as figure were longer when their fixation/attention was located on that figure than when it was located on the other figure, regardless of the figure's prototypicality. The intention effects were larger for prototypical shapes than for nonprototypical shapes, and for fixated/attended shapes rather than for unfixated/unattended shapes.

Discussion. These results suggest an role for spatial attention that is independent of structural knowledge. Spatial attention may operate through location directed, content-free mehanisms (LaBerge & Brown, 1989; Peterson & Gibson, 1989 (see appendix)). But, before these effects can be attributed to spatial attention as opposed to fixation, fixation and attention must be separated. This has yet to be done with the figure-ground stimuli, but, in a series of experiments using a partially biased Necker cube, Brad Gibson and I separated fixation, spatial attention, and intention. I summarize these experiments next.

Experiment 5 - 7

In three experiments, we examined whether spatial attention could be directed to parts of objects, or whether spatial attention to an entire object was obligatory, as Kahneman and Henik (1977; 1981) and Kahneman and Treisman (1984) have claimed. We used the opposed-set procedure with partially biased Necker cubes like the one shown in Figure 3. Intersection 1 in Figure 3 is biased toward the interpretation that the cube is facing downwards and to the left, whereas intersection 2 remains unbiased. If responses to all attributes of the attended object are obligatory, as Kahneman and his colleaques claim, the cube should always be seen in the downwards-facing orientation. In other experiments in which fixation and attention have not been separated, Peterson and Hochberg (1983; Hochberg & Peterson, 1987) have shown that observers can perceive the interpretation inconsistent with the bias when they fixate on intersection 2, suggesting that Kahneman and Treisman's proposal does not extend to objects constructed from structurally relavant parts and surfaces (or to drawings of such objects). Because fixation and

attention covaried in the previous experiments, their results may have been due to discriminability or to hemifield specific effects (cf. Peterson & Gibson, 1989, appendix).

Accordingly, Gibson and I separated fixation and attention in order to examine whether spatial attention could be focused on object parts, and as a result, could influence the perceived organization fitted to the object. In addition, we varied the observers' perceptual intentions while holding their spatial attention location constant to see whether any residual effects of intention remained when spatial attention was held constant. In our experiments, observers fixated one of three fixation points, one to the left of the cube, one to the right of the cube, and one in the center of the cube. While fixating each of these fixation points, observers were instructed to attend to either the top intersection or the bottom intersection of the cube (1 and 2, respectively, in Figure 3) and while holding their attention there, to try to hold either the interpretation consistent with the biased intersection or the interpretation inconsistent with that bias.

<u>Results</u>. We found that, when instructed to do so, observers were able to restrict their attention to the unbiased intersection and to perceive the interpretation inconsistent with the biased intersection, even when their fixation point was closer to the biased intersection than to the unbiased intersection. An experiment in which eye movements were monitored showed that these effects were not due to differences in eye movements (see Exp. 3 in Peterson & Gibson, 1989, in appendix).

<u>Discussion</u>.

Experiments 5 - 7 showed that spatial attention can be directed to parts of objects, and as a consequence, can influence the perceptual organization fitted to the entire object. Perceptual intentions do not operate simply through spatial attention: when spatial attention was held constant at the unbiased intersection, viewers were able to hold both interpretations for the cube. On the basis of the results of these experiments, we proposed the <u>functional equivalence model of perceptual organization</u>. We argue that spatial attention operates on content-free location representations (cf. LaBerge & Brown, 1989). When spatial attention is directed to a part of an object, the processing of the stimulus information in the location to which attention is directed is facilitated (i.e., the processing is speeded, so that the amount of stimulus information or detail processed in a given time period is increased) and the processing of the information in unattended locations is

inhibited (i.e., the processing of information in the unattended region is slowed or attentuated (Treisman, 1960)). The different patterns of faciliation and inhibition obtained when attention is directed to different locations result in functionally different structural descriptions of the cubes. For example, when attention is directed to the biased intersection of a partially-baised cube, the processing of the information there is facilitated, thereby speeding the processing of the occlusion and shading details. With this depth information strongly present in the structural description, and the information at the unbiased intersection attenuated, the structural description of the partially biased Necker cube may be functionally equivalent to that of a more densely biased cube, in that the best matching object representation will be the same as that accessed by a more densely biased cube. On the other hand, when attention is directed to the unbiased intersection, the combined facilitation of the processing of the unbiased intersection and inhibition of the processing of the biased intersection results in a structural description that is functionally equivalent to that of an unbiased cube, in that the set of potentially matching object representations is equivalent to the set accessed by a fully ambiguous cube.

The model that begins to emerge on the basis of these experiments is that spatial attention facilitates the processing of attended locations and inhibits the processing of unattended locations, regardless of their contents. Perceptual intentions can operate through the allocation of spatial attention (see Exp. 4), but intentions also operate through priming of structural knowledge regarding the intended interpretation. (ss Exps. 1 - 3).

Remaining issues.

The dual nature of selectivity in perceptual organization must be explored further. We are currently conducting a version of Experiment 4 in which fixation and attention are separated to be certain that the effects obtained in Experiment 4 were due to attention and not to fixation. The next step will be to examine whether the two types of selectivity interact under any conditions, or whether spatial attention is necessarily content-free.

III. Imagery.

Other experiments conducted during my first year of joint NSF/AFOSR support examined the role of structural knowledge in reversals of images. Chambers and Reisberg (1985) claimed that mental images were not reversible. In their view, different

interpretations cannot be fit to a single image. They argue that when so called "reversals" of images occur, one image is simply replaced by another. Of course, only a viewer who knows about the two alternatives can replace the original image with another image. In support of this view, they presented evidence that naive observes could not reverse the duck-rabbit in imagery.

Other research has shown that images can reverse, however. Finke, Farah, and Pinker (1989) showed that subjects can conjoin letters of the alphabet and discover new shapes that emerge from the conjunction. In my grant proposal, I argued that the difference between the two sets of results could be attributed to structural knowledge regarding the parts from which the emergent shapes were constructed. That is, Finke at al's subjects knew both the components from which the shapes were to be constructed and the relative locations of the components, whereas Chambers and Reisberg's subjects did not. In a series of experiments conducted with John Kihlstrom, Pat Rose, and Martha Glisky, I examined whether structural knowledge facilitated reversals in imagery. (As part of this series of experiments, we attempted to distinguish between two attributes of imagery: vividness and control. See Kihlstrom,. Glisky, Peterson, Harvey, & Rose, 1989, in appendix.

Experiments 8 - 10

We used three groups of subjects in each experiment. One group was shown the whole figure and asked to form a mental image of it. The other two groups were shown the duck-rabbit stimulus in three parts, each shown individually, and were asked to construct a mental image from the parts. One group was shown "natural parts", that is the stimulus was partitioned at minima of curvature. The other group received the parts parsed into "unnatural" parts, that is, parsed at places other than the minima of curvature. We reasoned that presenting the stimulus parsed into unnatural parts would interfere with the reversibility of the stimulus.

Results. Our prediction regarding the natural and unnatural parts was confirmed: Subjects given the "natural" parts reported significantly more reversals than subject giver the "unnatural" parts. However, Chambers and Reisberg's original results were also confirmed: We found that unless subjects were given either explicit or implicit instructions to relabel the back and the front of the shape, reversals were not obtained. In one experiment, this strategy was suggested implicitly by the inclusion of the goose-bird (which reverses in the same way as the duck-rabbit) as a training stimulus. In two other experiments, the instruction to consider the back of the head of

one animal as the front of the head of some other animal served as our strategic manipulation. When this strategy was suggested, both viewers who were shown the full figure and viewers who constructed the full figure from the natural parts reported reversals, whereas those who constructed the figure from unnatural parts reported no reversals.

These results suggest that, like the structural representations primed by the intention instructions, images may carry a reference frame that can be changed only with strategic intervention. Although changes in the coordinate axes may require strategic intervention, changes in the meaning of individual parts may not. That is, reversals requiring only a change in the meaning of individual parts of the shape, with no re-labeling of front and back or up and down may be possible in imagery. An experiment currently in progress is exploring that prediction.

Remaining issues.

Further research directed toward explicating structural knowledge effects in both imagery and perception will surely be useful in elucidating the nature and function of structural knowledge about shapes, and in particular, the role of reference frames and orientation in shape representations.

References

Biederman, I. (1987). Recognition by components: A theory of human image understanding. Psychological Review, 94, 115-147.

Chambers, D. & Reisberg, D. (1985). Can mental images be ambiguous? Journal of Experimental Psychology: Human Perception and Performance, 11, 317-328.

Corballis, M.C. (1988). Recognition of disoriented shapes. Psychological Review, 95, 115-123.

Finke, R.A., Pinker, S. & Farah, M.J. (1989). Reinterpreting visual patterns in mental imagery. Cognitive Science, 13, 51-78.

Hochberg, J. & Peterson, M.A. (1987). Piecemeal Perception and cognitive components in object perception: Perceptually coupled responses to moving objects. Journal of Experimental Psychology: General, 116, 370-380.

Hoffman, D.D. & Richards, W.A. (1984). Parts of recognition. Cognition, 18, 65-96.

Jolicouer, P. (1985). The time to name disoriented figures. Memory and Cognition, 13, 289-303.

Kahneman, D. & Henik, A. (1977). Effects of visual grouping on immediate recall and selective attention. In S. Dornic (Ed.), Attention and performance, VI. Hillsdale: LEA.

Kahneman, D. & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J.R. Pomerantz (Eds.), Perceptual organization. Hillsdale: LEA.

Kahneman, D. & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D.R. Davies (Eds.), Varieties of attention. New York: Academic Press.

Koffka, K. (1935). Principles of gestalt psychology. New York: Harcourt, Brace.

LaBerge, D. & Brown, V. (1989). Theory of attentional operations in shape identification. Psychological Review, 96, 101-124.

Marr, D. & Nishihara, H.K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. Proceedings of the Royal Society of London, B200, 269-294.

McClelland, J.L. & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. Psychological Review, 88, 375-407.

Palmer, S.E. (1975). Visual perception and world knowledge: Notes on a model of sensory-cognitive interaction. In D.A. Norman & D.E. Rumelhart (Eds.), Explorations in cognition. San Francisco

Peterson, M.A. (1984). Measures of selective components in perceptual organization. Unpublished doctoral dissertation, Columbia University, New York.

Peterson, M.A. (1986). Illusory concomitant motion in ambiguous stereograms: Evidence for nonstimulus contributions to perceptual organization. Journal of Experimental Psychology: Human Perception and Performance, 12, 50-60.

Peterson, M.A. & Hochberg, J. (1983). Opposed-set measurement procedure: A quantitative analysis of the role of local cues and intention in form perception. Journal of Experimental Psychology: Human Perception and Performance, 9, 183-193.

Riddoch, M.J. & Humphreys, G.W. (1987). Picture naming. In G.W. Humphreys & M.J. Riddoch (Eds.), Visual object processing: A cognitive neuropsychological approach. Hillsdale: LEA.

Roediger III, H.L. & Weldon, M.S. Reversing the picture superiority effect.

Treisman, A. (1960). Contextual cues in selective listening. Quarterly Journal of Experimental Psychology, 12, 242-248.

Ullman, S. (1984). Visual Routines. In S. Pinker (Ed.) Visual Cognition. Cambridge, MA: MIT Press.

Table 1
Change in Mean Duration of a Response
Part 2 - Part 1
(Seconds)

	Practice Control		Experimental "Knowledge"	
	Nonprototypical Center	Prototypical Surround	Nonprototypical Center	Prototypical Surround
I	-0.35	-0.30	-12.58	7.12
I-U	0.44	-0.11	-3.48	5.45

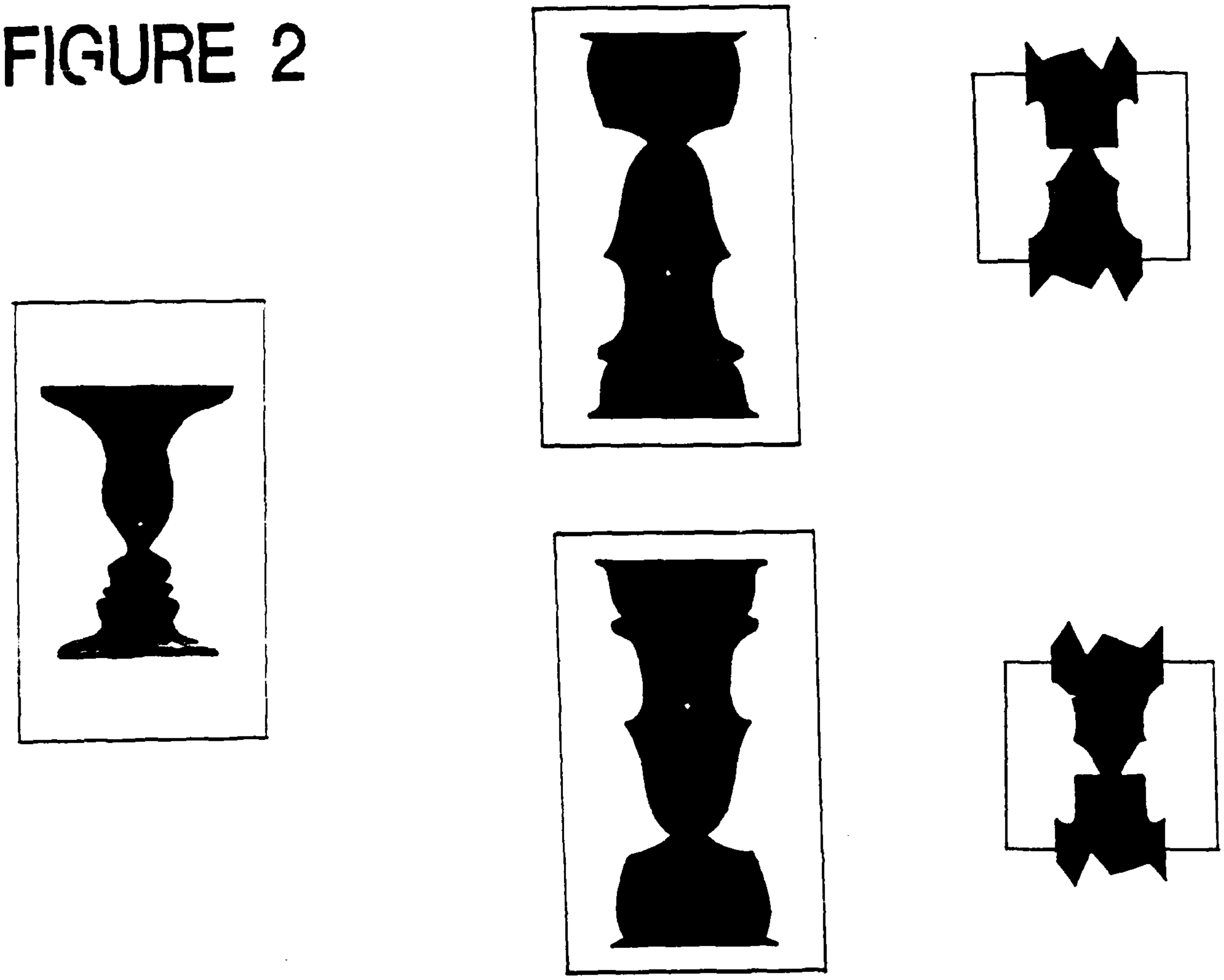
Table 2
Change in Mean Duraticn of a Response
Part 2 - Part 1
(Seconds)

	Practice Control		Experimental"Knowledge"	
	Nonprototypical Center	Prototypical Surround	Nonprototypical Center	Prototypical Surround
I	-2.32	-0.63	-2.79	0.49
I-U	-1.45	0.54	-1.14	0.09

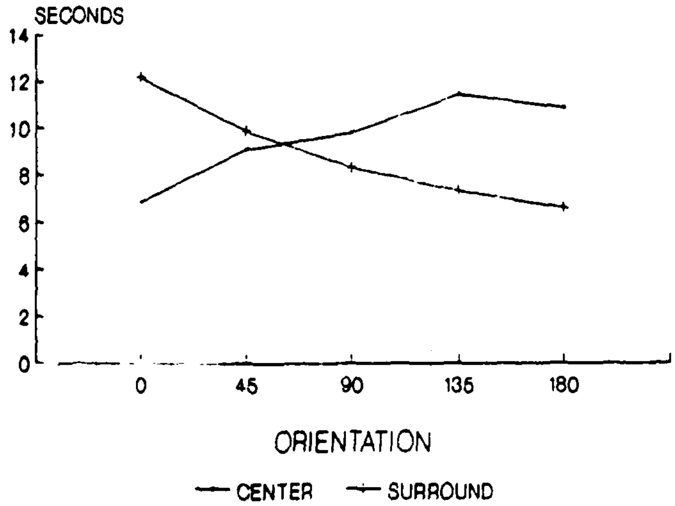
Table 3
Change in Mean Duration of a Response
Part 2 - Part 1
(Seconds)

	Practice Control		Experimental"Knowledge"	
	Nonprototypical Center	Prototypical Surround	Nonprototypical Center	Prototypical Surround
I	-2.59	-0.41	-1.73	0.15
I-U	-0.95	-0.48	-0.63	0.63

FIGURE 2

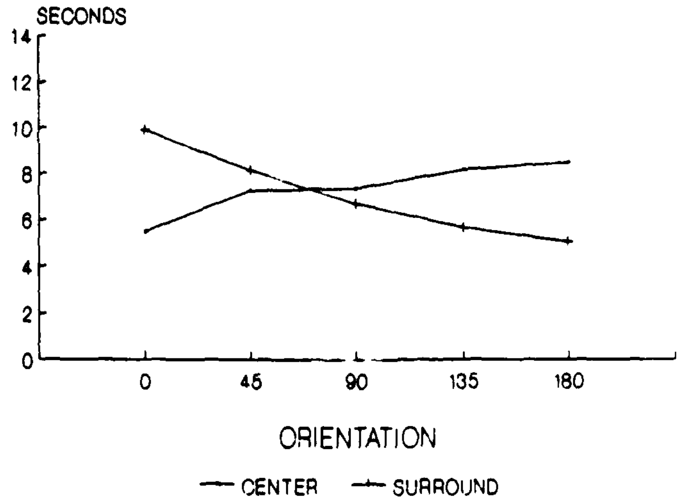


INTENDED



CENTER = CENTER SEEN AS FIGURE
SURROUND = SURROUND SEEN AS FIGURE

I - U



CENTER = CENTER SEEN AS FIGURE
SURROUND = SURROUND SEEN AS FIGURE

Presentations and Papers

Papers

Peterson, M. A., and Gibson, B. S. (1989). Directing spatial attention to parts of objects: The functional equivalence model of perceptual organization. (Manuscript submitted). (Copy in appendix.)

Kihlstrom, J. F., Glisky, M. L., Peterson, M. A., Harvey, E. M., & Rose, P. M. (1989). Vividness and control of mental imagery: A psychometric analysis. Journal of Mental Imagery, in press.

Posters and Presentations

Peterson, M. A. (1988, November). Shape generability: cognitive mediation in maintaining form perception. Poster presented at the Psychonomic Society meeting, Chicago, IL.

Peterson, M. A. (1989, April). Marr's methodology of vision ten years later: About reference frames and modularity. Paper presented at the meeting of the Society for Philosophy and Psychology, Tucson, AZ.

Peterson, M. A. (1989, June). Canonical descriptions, reference frames, and viewer intention in figure-ground organization. Paper presented at the American Psychological Society meeting, Alexandria, VA.

APPENDIX

Peterson, M. A., and Gibson, B. S. (1989). Directing spatial attention to parts of objects: The functional equivalence model of perceptual organization. (Manuscript submitted). (Copy in appendix.)

Kihlstrom, J. F., Glisky, M. L., Peterson, M. A., Harvey, E. M., & Rose, P. M. (1989). Vividness and control of mental imagery: A psychometric analysis. Journal of Mental Imagery, in press.

Peterson, M. A. (1989, June). Canonical descriptions, reference frames, and viewer intention in figure-ground organization. Paper presented at the American Psychological Society meeting, Alexandria, VA.

Directing Spatial Attention to Parts of Objects:

Functional Equivalence Model of Perceptual Organization

Mary A. Peterson and Bradley S. Gibson

University of Arizona

Running Head: ATTENDING TO PARTS OF OBJECTS

Correspondence:
Mary A. Peterson
Department of Psychology
University of Arizona
Tucson, AZ 85721

Abstract

Three experiments investigated whether spatial attention can be directed to parts of objects or whether responses to all parts or properties of an attended object are obligatory. Previous experiments have not been conclusive, either because fixation and attention locations were not separated or because the perceptual objects or responses used were unsatisfactory. We manipulated observers' fixation location, their spatial attention location, and their perceptual intentions to see one alternative of a partially-biased Necker cube, and measured their responses about the cube's perceived organization. The results show that spatial attention can be directed to parts of objects, thereby influencing the organization fitted to the object. On the basis of these experiments, we present a model of spatial attention and perceptual organization which argues that spatial attention operates through both facilitative and inhibitory mechanisms that determine the functional nature of the structural description of an object. We discuss distinctions between attention and intention and conclude that spatial attention can be directed to subregions of objects, whereas perceptual intentions may operate holistically.

Directing Attention to Parts of Objects:

Functional Equivalence Model of Perceptual Organization

This paper is directed to the question of whether it is possible to pay attention to certain parts of an object while ignoring others or whether attention must be directed to an object in its entirety. Kahneman and Treisman (1984) and their colleagues (Kahneman & Henik, 1977; 1981; Treisman, 1988) have argued that when attention is directed, it is directed to an entire object. In their view, attention is divided among all attributes, parts, and properties of the attended object. We refer to this hypothesis as the object-wide-attention hypothesis. On the other hand, Hochberg and Peterson (1987; Hochberg, 1968; 1982; Peterson & Hochberg, 1983) have argued that attention can be directed preferentially to parts of an object. They have shown that perceptual organization can depend more upon an object part to which fixation and attention are directed than upon unfixated, unattended object parts. We refer to this hypothesis as the piecemeal perception hypothesis. The tasks and objects upon which the evidence supporting these two hypotheses rests are very different. In this paper, we first evaluate the different claims and consider whether the differences can be attributed to different definitions of "attention" and "object." Next, we present experiments that show (a) that attention can be allocated to object parts and (b) that the allocation of attention to a part of an object can influence the perceptual organization fitted to that object.

Object-wide-attention hypothesis.

The object-wide-attention hypothesis arose in reaction to the typical paradigm used in spatial attention experiments: typical tasks are detection tasks and naming tasks; typical stimuli are brief light flashes or unconnected letters or shapes. Experiments using these tasks showed that when visual attention is focused on a location in space, the processing of items in that location is facilitated relative to items in other uncued locations, even if the uncued locations are closer to fixation (Downing, 1988; Downing & Pinker, 1985; Erikson; Erikson & Hoffman, 1972a&b; Posner, Snyder, & Davidson, 1980). In addition, spatial attention experiments yielded estimates that the focus of spatial attention could be as small as one degree of visual angle (Downing, 1988; Downing & Pinker, 1985; Erikson & Hoffman, 1972a&b; but see Merikle & Gorewich, 1982).

Kahneman and Henik (1977; 1981) argued that these estimates of the minimum size of attentional focus would not generalize to situations involving preattentively defined objects larger than one degree of visual angle. Their preattentively defined objects were constructed from items grouped together by variables like proximity and similarity, which may operate preattentively (Beck, 1972; 1982; Treisman, & Gelade, 1980). Kahneman and Henik proposed that attention can be directed to regions that have been identified as "objects" by preattentive processes but cannot be directed to subregions of such objects. Furthermore, Kahneman and Henik argued that once attention is directed to an object, responses to all attributes, parts, and properties of that object are obligatory (Kahneman &

Henik, 1981, p. 183; Kahneman & Treisman, 1984, p. 44 - 46). To test this hypothesis, Kahneman and Henik used a Stroop task and a memory task rather than the traditional spatial attention paradigm.

Stroop evidence. In the Stroop task, two shapes (e.g., a circle and a square) were exposed briefly, one on either side of a fixation point. Each shape contained either a non-color word or a color word printed in an ink color that was either the same as or different from the color denoted by the word. The observers' task on any given trial was to name the ink color of the word within either the circle or the square; this instructed shape was considered the attended shape. More Stroop interference and facilitation came from color words in attended shapes than from color words in unattended shapes, which Kahneman & Henik (1981) took as evidence that once attention is directed to an object, responses to all parts and properties of that object are obligatory.

We question whether Stroop effects can support this claim. Simply placing a word inside a shape may not make the word a part of a new emergent object. The word and shape remain structurally distinct, with the word serving as surface decoration on the shape rather than as a structural part of the shape. When one word is substituted for another, the lexical referent surely changes, but the perceptual object does not. For both empirical and theoretical reasons that are spelled out below, we question whether the results obtained with "objects" such as those used by Kahneman and Henik generalize to objects in which the to-be-ignored parts are structurally relevant to the attended object.

Furthermore, as Treisman (1969) has made clear, selective attention operates in different ways depending on the nature of the observer's task. For example, Treisman distinguished among three types of selection that might be mediated by attention: the selection of the _inputs_ to which to attend, the selection of which _analyzers_ to deploy on those inputs, and the selection of which _responses_ to make. Inasmuch as Stroop interference may result from a failure to inhibit a reading analyzer rather than from a failure to restrict spatial attention to a part of the attended shape as input, Stroop effects may not be relevant to questions regarding the minimum focus of spatial attention.

Memory evidence. Other evidence supporting the object-wide-attention hypothesis was obtained using a memory paradigm (Kahneman & Henik, 1977). The stimuli for the memory task were strings of unrelated letters and numbers grouped into subsets by proximity and color similarity. Observers were given brief exposures of these strings and were instructed to try to recall as many of the items as possible. Kahneman and Henik (1977) found that recall performance was uniformly high for items in the first (leftmost) of the two groups of letters or digits and uniformly low in the second (rightmost) group. Within these groups, all items were recalled with approximately equal probability, as long as there were no more than four items in the group. The average recall performance decreased as the number of items in the group increased.

Kahneman and Henik (1977) used this phenomenon to explore whether attention could be directed to parts of an object by examining whether

recall performance was affected by the addition of irrelevant "suffixes" to the group. Observers were instructed to ignore these suffixes, which differed from the other items in the group by virtue of their category membership (i.e., they were letters) but not by virtue of their color. Accuracy of recall for items in the groups containing the suffixes was similar to the accuracy of recall for items in groups containing an extra item, suggesting that observers' attempts to ignore these suffixes were unsuccessful.

We do not regard these memory results as unequivocal evidence for the object-wide-attention hypothesis for a number of reasons. First, the item strings containing an irrelevant suffix were presented randomly among item strings containing only digits, and they occurred on less than 20% of the trials. Observers were not informed about when a suffix trial would occur, so they could ignore the suffix only once they had recognized it. Nor were observers precued regarding the location in space which they were to ignore (or, conversely, the locations to which they were to attend). Thus, these experiments may not speak to the issue of whether spatial attention can be directed to subregions of preattentively grouped items.

Second, the possibility that the observers did not even attempt to restrict initial processing to the numbers must be considered. Research in the partial report tradition (also using memory tasks) shows that accuracy of recall is higher with selection by color than it is with selection by category membership (Sperling, 1960; von Wright, 1968). In particular, although observers can select items by category, it takes longer than

selection by color (Merikle, 1980). Accordingly, it is possible that observers participating in a memory task with these "objects" would not attempt to restrict analysis to the digits, whereas observers participating in other types of tasks with other types of objects might be able to successfully restrict attention to object parts.

Third, we question Kahneman and Henik's definition of "object" in the memory task as well as in the Stroop task. For the memory task, "objects" were created from strings of unrelated items by virtue of color similarity. But evidence obtained with objects contructed from unrelated items may not generalize to other objects in which the local structure (i.e.. the structure of the parts) is informative about the global structure of the object (Pomerantz, 1983). Furthermore, even though both similarity and proximity have been shown to group items together preattentively (Beck, 1972; 1982; Treisman & Gelade, 1980), such variables may not be relevant to object identification (e.g., see Biederman, 1987). Many processing steps are likely to occur after preattentive grouping and before the perceived interpretation (cf. Treisman, 1986). Consequently, the relationship between grouping processes and other organizing processes may not be a simple one. Moreover, empirical evidence from experiments using multistable objects constructed from structurally relevant parts and surfaces (and drawings of such objects) suggests that observers may be able to attend selectively to one part of an object while ignoring another part of the same object (Hochberg & Peterson, 1987; Peterson & Hochberg, 1983; 1989; Tsal & Kolbet, 1985; Olsen & Orhbach, 1966). We consider that research next.

Spatial attention and the organization of multistable figures.

Hypotheses about how allocating spatial attention to parts of an object might influence its perceived organization have typically rested on the constructivist assumption that the first glance at a figure instantiates an hypothesis about the identity of the figure and that subsequent glances are interpreted in the context of this hypothesis (Hochberg, 1968; Neisser, 1967). In this view, figural reversals occur because alternative hypotheses are tested and confirmed over a series of successive glances. An implicit assumption of this view is that some parts of a reversible figure must favor one interpretation whereas other parts must favor the alternative intepretation, although these distinctive parts must be capable of being assimilated into the alternative interpretations (Chastain & Burnham, 1975). A prediction that follows from this view is that the location of the eyes just prior to reversal should reveal the distinctive features for each interpretation.

Compelling emprical evidence in support of this hypothesis has not been obtained, however. Eye movements do seem to be coupled to reversals, but it is often unclear whether eye movements precede or follow reversal (Ellis & Stark, 1979; Pheiffer, Eure, & Hamilton, 1956). When there was reasonable certainty that eye movements occurred prior to reversal, the location to which the eyes moved did not vary convincingly with the size of the objects and hence, with the location of the distinctive features (Ellis & Stark, 1979). Moreover, it has been shown that reversals can occur in

stabilized images when eye movements cannot bring different parts of the object into foveal vision (Pritchard, 1958).

The failure to find a reliable relationship between eye location and reversal does not rule out the distinctive features hypothesis of reversal, however, because the direction of gaze does not necessarily reveal the location of spatial attention (Posner et al, 1980; Remington, 1980). Tsal and Kolbet (1985) altered the distinctive features hypothesis slightly by proposing that merely focusing attention on the distinctive feature for a given interpretation might cause a reversal into that interpretation or might lengthen the duration of perceiving that interpretation. This version of the distinctive features hypothesis clearly rests upon the assumption that attention can be preferentially directed to parts of a reversible figure. Nevertheless, Tsal and Kolbet's evidence that attention can be directed to parts of an object is not conclusive, as we discuss next.

Tsal and Kolbet (1985) used a brief exposure of the Jastrow duck/rabbit figure, and instructed their viewers to try to see the rabbit on some trials and to try to see the duck on other trials. On trials where they had been instructed to try to see the rabbit (duck), observers were faster to name letters flashed immediately after the offset of the multistable figure that were located on the side of fixation where the part judged as distinctive for the rabbit (duck) interpretation had been. These findings support the conclusion that spatial attention was located on different sides of fixation following brief exposures of multistable figures to which viewers were instructed to try to fit different

interpretations. However, since Tsal and Kolbet's subjects did not report about the perceived organization of the multistable figure, but only named the letters, it is impossible to determine whether the attentional movements were of any consequence for the response made to the multistable figure itself (i.e., to the perceived organization). Therefore, these experiments cannot be used to rule out the object-wide-attention hypothesis.

Other evidence suggesting that attention can be directed to less than an object has been obtained using paradigms in which viewers respond about the perceived organization of three-dimensional objects constructed from structurally relevant parts (or drawings of such objects) (Hochberg & Peterson, 1987; Peterson & Hochberg, 1983; 1989). The basic procedure, the opposed-set procedure (Peterson & Hochberg, 1983), was one in which viewers' perceptual intentions were manipulated while they fixated one of two intersections of a partially-biased Necker cube such as that shown in Figure 1. The intersections differed in that one was "biased" toward one potential interpretation of a Necker cube (see intersection 1 in Figure 1 where occlusion and shading specify that the horizontal line is in front of the vertical line, and hence, that the cube is facing downwards and to the left), whereas the other was "unbiased" (see intersection 2 in Figure 1 where no local information specifies that either the horizontal or the vertical line is in front.)

Insert Figure 1 about here

Observers were instructed both to fixate one of these intersections and to keep their attention around the fixated area during 30 second trials. In addition to manipulating fixation/attention instructions, Hochberg and Peterson manipulated observers' perceptual intentions through instructions to try to hold a local interpretation consistent with one of the two potential interpretations for the Necker cube (e.g., observers were instructed to try to keep either the horizontal or the vertical line appearing in front). Thus, Hochberg and Peterson manipulated two types of selective processing in the opposed-set paradigm: They manipulated the selection of input through the fixation/attention instructions, and they manipulated the selection of which interpretation to fit to the multistable figure through the intention instructions.

Previous experiments (Hochberg & Peterson, 1987; Peterson & Hochberg, 1983; 1989) showed that when observers fixated on and attended to the biased intersection, they were very successful at holding the interpretation consistent with the bias and quite unsuccessful at holding the interpretation inconsistent with the bias. Thus, the bias present at the biased intersection was effective in determining the perceived organization of the cube when fixation and attention were fixed on the biased intersection. However, when observers fixated on and attended to the unbiased intersection, which was less than two degrees of visual angle from

the biased intersection, the biased intersection had no observable effect on the perceived organization of the cubes: Viewers were approximately equally successful at holding both potential interpretations of the cube when they fixated on and attended to the unbiased intersection.

Unlike Tsal and Kolbet (1985), Hochberg and Peterson have shown that perceived organization <u>per se</u> varies with the intention instructions. They did so in a series of experiments in which viewers reported about variables that are perceptually, but not consciously, coupled to the perceived organization (Epstein, 1980; Gogel, 1980; Gogel & Tietz, 1974; Hochberg, 1974). (Examples of variables that are perceptually coupled to the perceived depth organization of a cube are the perceived direction of rotation of a randomly oscillating three-dimensional cube (Hochberg & Peterson, 1987) and the perception of illusory concomitant motion in a stationary three-dimensional depth reversed cube viewed by moving observers (Hochberg & Peterson, 1987; see also Peterson, 1986).) Since these variables covary with perceived depth, reports about these variables can serve as indirect reports about perceived depth. These indirect reports replicated the results obtained with direct reports about perceived organization. (Reports about the perceptually coupled variables also suggest that the perceived organization of the entire cube changes with changes in the organization perceived at the fixated/attended intersection. Particularly compelling are the perceptions that the entire cube reverses direction of rotation when the perceived depth of a moving cube reverses and that the entire cube moves when the perceived depth of a stationary

cube reverses.)

In other experiments, Peterson and Hochberg have shown that the intention effects do not simply reflect different eye movements made under the different hold instructions (i.e., the effects of fixation and intention seem to be separable) (Peterson, 1984; 1986; Peterson, Harvey, & Weidenbacher, 1989). Thus, the previous opposed-set experiments provide robust evidence that the fitting of an interpretation to these figures does not entail an obligatory response to the biased intersection when fixation and attention are directed to the unbiased intersection. This effect would not be predicted by an object-wide-attention hypothesis.

Despite the fact that the opposed-set experiments are not subject to the same criticisms as Tsal and Kolbet's experiments, they do not provide uncontestable evidence contrary to the object-wide attention hypothesis, for the following reasons. First, it can be questioned whether the previous opposed-set results are attributable to the differential allocation of spatial attention within the figures or to the relative discriminability of the biased intersection from the two different fixation points. Inasmuch as fixation location and spatial attention location always covaried in those experiments, the spatial attention effects cannot be separated from fixation effects.[1]

Second, the opposed-set previous experiments do not permit the separation of spatial attention and intention either: Intention may be operating through the differential allocation of attention in space. Specifically, the location of the biased intersection with respect to the

two visual hemifields covaried with fixation location in the previous opposed-set experiments. When fixation was located on the unbiased intersection, the biased and unbiased regions of the cube fell in different hemifields, whereas when fixation was located on the biased intersection, the biased region of the cube extended into both hemifields. Some spatial attention experiments suggest that there may be a distinct advantage for ignoring stimuli located in the opposite hemifield (Downing & Pinker, 1985; Hughes & Zimba, 1985). Accordingly, it is possible that observers fixating the unbiased intersection adopted a strategy of attending to one hemifield when instructed to try to hold the interpretation consistent with the bias at the biased intersection and of attending to the other hemifield when instructed to try to hold the alternative organization. This strategy was not available with fixation on the biased intersection, because in that condition the biased information fell in both hemifields. Given the possibility that observers used this strategy to follow the opposed-set instructions, it remains possible that attention may not be able to be directed to a part of an object that falls entirely within one hemifield.

The experiments

In the experiments reported here, we separated the three variables of fixation location, spatial attention location, and intention so that we could examine (a) whether spatial attention could be directed to parts of an object when fixation was held constant and (b) whether perceived organization could vary with the observer's intentions when both spatial attention location and fixation location were held constant. Three

conditions of fixation were used: fixation to the left, to the right, and in the center of the cube, as shown in Figure 2. At each fixation location, each observer participated in two conditions of attention: attention to the biased intersection and attention to the unbiased intersection. In addition, observers' perceptual intentions were varied while fixation location and spatial attention location were held constant through instructions to try to hold either the horizontal or the vertical line forward.

Predictions. In the condition in which observers attend to the biased intersection, we expect to replicate the results obtained in the previous opposed-set experiments when observers fixated the biased intersection. That is, observers should show a double pattern of (1) success at following instructions to try to hold the organization consistent with the bias and (2) relative failure at following instructions to try to hold the organization inconsistent with the bias. If this effect is obtained, we will be confident that the occlusion and shading present at the biased intersection serves as an effective bias when it is not fixated. We will use this effect to look for an obligatory response to the bias when attention is directed to the unbiased intersection. If, as the object-wide-attention hypothesis predicts, responses to the biased intersection are obligatory, then the same double pattern of success at holding the consistent interpretation and relative failure at holding the inconsistent organization should be obtained when observers attempt to restict their attention to the unbiased intersection. Success at holding the consistent

interpretation when attention is directed to the unbiased intersection is not sufficient; the double pattern is necessary to conclude that responding to the biased intersection is obligatory. It is only the latter effect that truly reflects an obligatory response to the bias, the former effect could reflect a voluntary response to the bias.

If spatial attention can be directed to parts of an object when fixation and attention are separated, then, even if the bias is effective on trials on which observers attend to the biased intersection, no obligatory effect of the bias will be obtained on trials on which observers attend to the unbiased intersection: Observers should be able to hold both the interpretation consistent with the bias and the interpretation inconsistent with the bias when their attention is directed to the unbiased intersection (as they were when fixation and attention were both directed there in previous opposed-set experiments).

If the object-wide-attention hypothesis is correct only for objects that fall entirely within one hemifield, but not for objects that subtend both hemifields, we should obtain an effect of fixation location: The influence of the biased intersection should extend to the unbiased intersection when fixation is located to the left or to the right of the cube, but not when it is located in the center of the cube, because in the latter condition, the biased and unbiased intersections fall in opposite hemifields.

If the previous piecemeal perception results were due to the relative discriminability of the biased intersection from the two fixation points

and not to attentional allocation, we should obtain a different type of fixation effect: With center fixation, the biased intersection should exert an equal influence on the organization perceived at both intersections because in this condition, the biased and unbiased intersections are equally far from fixation and hence, should be equally discriminable. The influence of the biased intersection should be greatest in the condition where fixation is closest to the biased intersection (the right fixation in Figure 2) and should be least for the condition where fixation is farthest from fixation (left fixation in Figure 2).

The three experiments reported here show that attention can be directed to parts of an object when fixation and attention are separated. In Experiments 1 and 2, fixation instructions were used and eye movements were not monitored. Experiments 1 and 2 differ by the addition of an unbiased control cube to Experiment 2 which is used to provide a baseline measure of reversals under the conditions used in these experiments. In Experiment 3, we monitored eye movements so that we could be more certain that observers were not moving their eyes differently in the different conditions of instructed intention.

Experiment 1

In this experiment, we used the opposed-set procedure with the two cubes shown in Figure 2.

Insert Figure 2 about here

Both cubes are biased at the bottom intersection; The figure shown in the top row of Figure 2 is biased by shading and occlusion toward the orientation that faces downwards and to the left, and the figure shown in the bottom row of Figure 2 is biased toward the orientation that faces upwards and to the right. If observers can restrict their spatial attention to one intersection of these cubes so that when attention is directed to the unbiased intersection the organization inconsistent with the biased intersection can be perceived, that would suggest that the object-wide-attention hypothesis is incorrect.

Method

Subjects. Subjects were 12 students at the State University of New York at Stony Brook who participated in this experiment in order to fulfill a requirement for an introductory psychology course. All subjects had vision that was normal or corrected to normal.

Stimulus and Apparatus. The stimuli were line drawings of the cubes shown in Figure 2, positioned 84 cm from the subjects so that the cubes subtended 1.6 degrees of visual angle and the distance between the biased and unbiased intersections was approximately 1 degree of visual angle. From the left and right fixation points, the distance to the nearest intersection was approximately .8 deg, the distance to the farthest intersection was approximately 1.4 deg, and the distance to the farthest edge of the of the object was 1.8 degrees. From the center fixation point, the distance to each of the two intersections was approximately .5 deg. The figures shown to the observers contained only the fixation point relevant for the current trial.

Subjects pressed keys to report whether the horizontal or vertical line appeared forward. The frequency, order, and duration of these key presses were recorded by a Rogers A6 Timer/Driver in an Apple IIe.

Procedure. Subjects participated in the experiment individually. They were shown a reversible cube and the two alternative interpretations were pointed out to them. They were told that we were examining the degree to which the viewer's intentions could influence how attended but unfixated objects were seen.

Before the experimental trials, observers participated in eight practice trials. On four of these trials they viewed an unbiased Necker cube (Figure 5c) from a center fixation point. On the other four practice trials, they viewed this unbiased cube from either the left or the right fixation point. (Half of the subjects viewed the cube from the left; the other half viewed the cube from the right.) These sets of four practice trials consisted of one trial with each hold instruction paired with instructions to attend to each of the "top" and "bottom" intersections. The hold instructions requested that viewers try to hold either the horizontal or the vertical line in front at the attended intersection. For Figures 2a - c, the instructions to hold the vertical line forward at the biased intersection and the horizontal line forward at the unbiased intersection were instructions to hold the organization consistent with the bias. For Figures 2d - f, the instructions to hold the horizontal line forward at the biased intersection and the vertical line forward at the unbiased intersections were instructions to hold the organization that was consistent with the bias. During each trial, viewers

pressed buttons to indicate whether the horizontal or the vertical line appeared forward. They pressed neither key when the attended intersection appeard flat or when their attention wandered from the instructed intersection.

After each of the practice trials, the importance of maintaining fixation on the fixation point and keeping their attention right around the location of the instructed intersection was stressed, as was the importance of signaling the occurrence of any reversal (regardless of how brief). Viewers were also reminded that they were to release both keys when they could not see what was happening at the instructed intersection or when their attention was not centered on that intersection.

Observers participated in 24 30s experimental trials. For each of three fixation points, there was one trial attending to each intersection while following each hold instruction. Observers viewed the experimental figures only during the experimental trials. They participated in all four trials with each figure-fixation point combination before viewing the next figure. Figure order, fixation order, and hold instruction order were counterbalanced within and across subjects.

Data analysis: Intention. For each 30 s trial, the total duration that viewers reported seeing the instructed line (I_t) in front and the total duration that they reported seeing the uninstructed line (U_t) in front were computed. The data from each trial were then represented as an intention index (Peterson, 1986),

$$I = (I_t - U_t) / (I_t + U_t).$$

When there is a substantial amount of time during the 30 s trials during which subjects press neither key, the intention index provides a good measure of the relative durations of seeing the two alternative interpretations. We chose to use this measure because we encouraged the observers to report lapses of attention or fixation by removing their fingers from both keys.[2]

I can range from +1 to -1 for a single condition of hold instruction and attention to an intersection, reflecting total success (+1) or total failure (-1) at following the intention instructions. We performed an analysis of variance (ANOVA) on the intention idices with four within-subjects variables (fixation location (left, center, or right); cube type (facing downwards or upwards); spatial attention location (unbiased intersection or biased intersection); and hold instruction (hold the interpretation consistent with the bias, or hold the interpretation inconsistent with the bias)).

In addition to the individual *I*s calculated for each intersection for each hold instruction, we examined the mean *I* across hold instructions for each intersection as a measure of the effectiveness of the viewer's intentions. If the intersection is malleable (i.e., subject to influence from the viewers' intentions), the mean *I* will be significantly greater than zero. If the intersection is not malleable (i.e., if only one organization is seen regardless of the viewer's intentions or if the cube alternates between two percepts regardless of the viewer's intentions), the mean *I* for a given intersection will not differ from zero (see Peterson, 1986).

Lapses. We summed the durations that neither key was pressed during each trial as a measure of lapses in following the fixation and attention

instructions. This measure can serve as a secondary measure of the effects of the biased intersection. For example, if it is difficult to hold the instructed interpretation, viewers might allow their attention to wander from the instructed intersection in order to increase the effectiveness of their intentions. This might happen more often on trials on which observers try to hold the inconsistent organization than on trials on which they try to hold the consistent organization. If it does, then lapses can serve as a measure of the difficulty of restricting attention within an objects.

Results

Intention effects The ANOVA on the individual *I*s failed to show any effects of fixation: The main effect of fixation location was not significant, $F < 1$, nor were any interactions involving fixation location, all *p*s > .35, thus ruling out both the discriminability and the hemifield location explanations of the previous opposed-set results.

Observers were more successful at holding the interpretation consistent with the bias than the interpretation inconsistent with the bias, $F(1, 11) = 5.2$, $MS_e = .25$, $p = .044$. As can be seen in Figure 3, however, the difference between the ability to hold the consistent and the inconsistent intepretations was greater at the biased intersection than at the unbiased intersection, as indicated by an interaction between intersection and hold instruction, $F(1, 11) = 9.31$, $MS_e = .38$, $p = .011$.

Insert Figure 3 about here

A marginal three-way interaction among cube-type, attended intersection, and hold instruction, $F(1, 11) = 4.47$, $MS_e = .05$, $p = .058$, showed that the Is obtained at the unbiased intersection for hold consistent and hold inconsistent trials differed for the downwards facing cube only, $t(11) = 2.74$, $p < .02$, but not for the upwards-facing cube, $t < 1$.

The I obtained at the biased intersection for hold inconsistent trials was unusually large ($M = .36$). Indeed, both intersections of the cube were malleable: The mean Is across hold instructions for both the biased and the unbiased intersections were greater than zero for both cubes, all ps $< .01$, indicating that the organization perceived at both intersections of the cube was influenced by the viewer's intentions.

Lapses. The observers in Experiment 1 indicated lapses in fixation or attention for an average of 5.31 s per trial. An ANOVA showed that observers reported more lapses when they tried to hold the inconsistent interpretation, $F(1, 11) = 9.88$, $MS_e = 6.80$, $p = .009$, although an interaction between intersection and hold instruction, $F(1, 11) = 7.67$, $MS_e = 11.17$, $p = .018$, shown in Figure 4, indicated that this was true only when with attention directed to the biased intersection, $t(11) = 3.15$, $p < .01$; with attention directed to the unbiased intersection, the durations of lapse time reported for both hold instructions were the same, $t < 1$.

Insert Figure 4 about here

In addition, an interaction between cube-type and hold instruction, $F(1, 11) =$

10.18, $MS_e = 6.15$, $p = .009$, showed that the lapses reported for the hold consistent and hold inconsistent trials differed significantly in the downwards facing cube but not in the upwards facing cube.

Discussion

Experiment 1 did not provide a clear answer to the question of whether attention can be restricted to parts of an object. As would be predicted by the object-wide-attention hypothesis, observers attending to both the biased and unbiased intersections were more successful at holding the interpretation consistent with the biased intersection. On the other hand, the difference between the durations of seeing the consistent and inconsistent interpretations was larger at the biased intersection than at the unbiased intersection, suggesting that the response to the biased intersection may not have been obligatory. Indeed, for the upwards facing cube, there was no difference in the observers' ability to hold the consistent and the inconsistent interpretations when they attended to the unbiased intersection.

The lapse time measures do not support an object-wide-attention hypothesis. Although observers reported more lapses on hold inconsistent trials when they attended to the biased intersection, they reported equal lapses on hold consistent and hold inconsistent trials when they attended to the unbiased intersection.

None of the effects of location was significant, thereby clearly showing that the previous opposed-set results obtained by Hochberg and Peterson are not attributable to the location of the biased and unbiased intersections relative to the two hemifields, nor to the relative discriminability of the

two intersections from the different fixation points. Although the absence of fixation effects rules out a strict discriminability account, the fact that the biased intersection was less discriminable when unfixated in this experiment than when fixated in previous experiments might account for the fact that the _I_ obtained when observers attended to the biased intersection and tried to hold the inconsistent interpretation was larger than the _I_s typically obtained in that situation (see Hochberg & Peterson, 1987). This suggests that the effectiveness of the occlusion and shading information at the biased intersection was diminished when it was located at a distance of at least .5 degrees from fixation and, at least for the purposes of this task, the bias did not seem to be further diminished when it was removed by as much as 1.4 degrees from fixation.

Before we conclude that the bias is ineffective when it is not fixated, however, we should assess its effectiveness by comparing the reversal rate at the biased intersection to the reversal rate at the analogous intersection of an unfixated unbiased cube. Inasmuch as we do not know what the baseline reversal rate of an unbiased cube is when viewed under the conditions used in Experiment 1, we cannot be sure of the extent of the influence exerted either by the shading and occlusion at the biased intersection or by the viewers' intentions. Accordingly, an unbiased Necker cube was included as a control cube in Experiment 2 so that we could evaluate the effectiveness of the unfixated bias.

An unbiased cube can serve as a baseline for viewers' _a priori_ preferences as well. In Experiment 1, the _I_ obtained at the unbiased

intersection for the hold consistent trials was greater than that obtained for the hold inconsistent trials only for the downwards facing cube. This difference might reflect an influence from the biased intersection, but it might reflect other influences instead. For example, a larger I for hold consistent trials than hold inconsistent trials on the downwards facing cube might reflect the preference that viewers have for seeing a Necker cube in the orientation that faces downwards and to the left (e.g., Price, 1967). Alternatively, a larger I on hold consistent trials than hold inconsistent trials at the unbiased intersection might reflect a strategic widening of attention to include the biased intersection on hold consistent trials. Of course, if the increase in I on hold consistent trials at the unbiased intersection reflects a non-obligatory strategy adopted by the viewers, similar results would be expected for the upwards facing cube, unless *a priori* preferences to see the cube in the downwards facing orientation interacted with this strategy. Hence, including an unbiased cube as a baseline condition in Experiment 2 may eliminate the difference between the two cubes observed in Experiment 1.

Experiment 2

The stimuli used in Experiment 2 are shown in Figure 5. In addition to the unbiased Necker cube, we used cubes that were biased at the top intersection rather than at the bottom intersection to examine the generality of the effects obtained in Experiment 1.

Insert Figure 5 about here

Method

Subjects. The subjects were 12 students from the University of Arizona who participated in this experiment as part of a course requirement. All had vision that was normal or corrected to normal.

Stimulus materials, apparatus, and procedure. The viewing distance and stimulus dimensions were the same as those used in Experiment 1. Observers responses were collected and analyzed on a Compaq 386 microcomputer. The procedure was the same as that used in Experiment 1.

Data analysis. The responses to the reversible Necker cube served as a baseline against which to examine the responses to the two biased experimental cubes. As in Experiment 1, an intention index, I, was computed for each fixation point for each hold instruction for each intersection of each cube. Next, a bias influence index (B) was calculated for each of these conditions by subtracting the analogous I obtained for the unbiased control Necker cube (I_c) from the I obtained for the biased experimental cube (I_e),

$$B = I_e - I_c.$$

Thus, in Experiment 2, we measure the reversibility of the two intersections of the biased cube relative to the reversibility of the two intersections of the unbiased cube under similar conditions of fixation and intention. B can range from +2 to -2. A B of zero indicates that performance with the biased

cube did not differ from performance with the unbiased cube. If *B* is both significantly greater than zero when viewers attend to the biased intersection and try to hold the consistent organization and signficantly less than zero when they try to hold the inconsistent organization there, then the unfixated biased intersection will be considered effective. Responses to the biased intersection will be considered obligatory if *B* is both significantly greater than zero when viewers attend to the unbiased intersection and try to hold the consistent organization and signficantly less than zero when they try to hold the inconsistent organization there.

We also recorded the lapses in following the attention and fixation instructions (i.e., the durations that neither key was pressed) for both the biased experimental cube (L_e) and the unbiased control cubes (L_c). The lapses reported below are the differences between the lapses reported with the biased cube and the unbiased cube (L_e - L_c) under similar conditions of fixation, attention, and hold instruction.

Results

Bias influence index. The results of Experiment 2 are very similar to those of Experiment 1. The main effect of fixation was not significant, $F < 1$, nor was the main effect of cube-type, $F < 1$. In Experiment 2, the the main effect of hold instruction did not reach significance, $F(1, 11) = 3.19$, $MS_e = .06$, $p = .10$.

As can be seen in Figure 6, the interaction between intersection and hold instruction was significant, $F(1, 11) = 26.93$, $MS_e = .10$, $p = .0003$: The difference between the observers' ability to hold the interpretation

consistent with the bias and the interpretation inconsistent with the bias was greater at the biased intersection than at the unbiased intersection.

Insert Figure 6 about here

The bias present at the biased intersection was effective: _B_ was both significantly less than zero when observers tried to hold the inconsistent interpretation while attending to the biased intersection (M = -.14), $t(11) = -2.56$, $p < .05$, and significantly greater than zero when observers tried to hold the consistent interpretation while attending to the biased intersection (M = .14), $t(11) = 3.74$, $p < .005$. The biased intersection did not exert an obligatory influence at the unbiased intersection, however: _B_ was not significantly less than zero when observers tried to hold the organization inconsistent with the biased intersection, $t(11) = -1.57$, $p > .10$, although it was significantly greater than zero when observers tried to hold the organization consistent with the biased intersection, $t(11) = 2.99$, $p < .02$. No other effects in this ANOVA were significant.

Intention. A _B_ that does not differ from zero should not be taken as evidence that intention was not effective. For both the experimental and the control cubes, the _I_s obtained when viewers' attended to the unbiased intersection were greater than zero ($p < .001$) for both hold instructions.

Lapses. The mean lapse time reported in this experiment was 1.53 s for the biased experimental cubes (L_e) and 1.37 s for the unbiased control cubes (L_c). $L_e - L_c$ was greater when observers tried to hold the inconsistent

organization (.37 s) than the consistent organization (-.05 s), $F(1, 11) = 5.80$, $MS_e = 2.24$, $p = .035$. The means, although different, are very small, and a marginal three-way interaction among attended intersection, fixation, and hold instruction, $F(1,11) = 3.22$, $MS_e = 1.83$. $p = .06$, shown in Figure 7, indicated that although $L_e - L_c$ was greater than zero for observers trying to hold the inconsistent interpretation while attending to the biased intersection from all fixation points, it was only greater than zero for observers trying to hold the inconsistent interpretation while attending to the unbiased intersection from the right fixation point; from both the left and center fixation points. $L_e - L_c$ was *less* than zero for observers trying to hold the inconsistent interpretation while attending to the unbiased intersection.

Insert Figure 7 about here

Discussion

Experiment 2 provides evidence that the shading and occlusion at the biased intersection of the cube effectively bias the organization that can be perceived when attention is directed there. The *B* index revealed that observers were significantly more successful at holding the consistent interpretation and significantly less successful at holding the inconsistent interpretation when attending to the biased intersection of the biased cube than to the analogous intersection of an unbiased cube.

Responses to the biased intersection were not obligatory when observers

attended to the unbiased intersection, however: Even though *B* was greater than zero on hold consistent trials, it was not less than zero on hold inconsistent trials. An obligatory repsonse to the biased intersection should entail a decrease in observers' ability to hold the inconsistent interpretation at the unbiased intersection as well as an increase in their ability to hold the consistent interpretation. As long as there is no difference between observers' ability to hold the inconsistent interpretation of the experimental and control cubes when their attention is directed to the unbiased intersection, any increased success at holding the consistent organization of the experimental cube relative to the control cube in similar conditions of attention might be attributable to viewer strategy. The fact that we did not obtain a three way interaction among cube-type, intersection, and hold instruction in Experiment 2 is consistent with a strategy account (i.e., we would expect that viewers would use the same strategies on both cubes, ruling out a three way interaction). In Experiment 1, this effect may have been obscured because we did not have a baseline measure of observers' *a priori* abilities to hold the two interpretations on an unbiased cube.

Because of the inclusion of the control cube, Experiment 2 provides strong support that attention can be directed to parts of an object. The shading and occlusion effectively biased the perceived organization of both cubes when observers attended to the biased intersection, but the bias did not exert an obligatory effect when observers attended to the unbiased intersection of either cube. Furthermore, Experiment 2 shows that the differential allocation of attention within a reversible object influences the

perceived organization of that object.

The lapses of attention or fixation reported by the observers in Experiments 1 and 2 did not vary systematically with hold instruction or intersection, suggesting that viewers did not move their eyes or their attention differently under the different hold instructions. This finding is consistent with other experiments examining eye movements in the opposed-set procedure that have failed to find any differences in eye movements under the different instructions (Peterson, 1984; 1986). Given the importance of the separation of fixation and attention in these experiments, however, we wanted to be more certain that viewers did not move their eyes differently or center their attention on different locations under different hold instructions. Accordingly, we conducted an additional experiment in which we monitored eye and attentional movements more closely.

Experiment 3

In Experiment 3, we converted the left and right fixation points into circular fixation regions with diameter of approximately .5 degrees of visual angle and instructed observers to maintain a .25 degree after-image within the boundaries of the circular fixation region. We used this technique to monitor eye movements because the precision of measurement offered by the small perimeter of the fixation regions was greater than that offered by the recording devices typically used in the spatial attention literature (e.g., electrooculogram recordings, cf Dowing & Pinker, 1985). In addition, viewers' reports regarding the location of an after-image have served as good indicators of eye location in a variety of tasks (Alpern, 1971; Post &

Leibowitz, 1982). In Experiment 3, we examined whether the results obtained in Experiment 2 were replicated when an eye movement monitoring task was combined with the spatial attention task and the opposed-set task.

Method

Subjects. The observers were three psychology graduate students at the University of Arizona. Two were unaware of the experimental hypothesis. The third observer was the second author (BG) who was not aware of the specific hypotheses of this study at the time of testing. All observers had vision that was normal or corrected to normal.

Stimuli and apparatus. The bottom-biased downwards-facing cube (Figure 2a &c) and the unbiased Necker cube (Figure 5c) were used as stimuli in this experiment. They were viewed from the same distance as the cubes in the previous two experiments.

The solid fixation points employed in Experiments 1 and 2 were replaced by black circular outlines, measuring 0.7 cm in diameter, which subtended approximately .5 degrees of visual angle. Only two fixation areas were used in this experiment, one to the left of the cube and the other to the right of the cube. The circular frames were drawn so that the arc of the circle nearest the cube was located at the same distance from the cube's edge as the fixation points had been in the previous two experiments.

After-images were created using a Vivitar electronic flash attachment which was covered by an opaque black shield except for a small hole (0.2 cm in diameter). The flash attachment was mounted at a distance of 44 cm from the subjects, so that the after-image subtended approximately 0.25 degrees. A

2.5Hz strobe light was used to create a flickering field which has been shown to extend the duration that an after-image is visible (e.g., Hochberg & Hay, 1956).

In other respects, the stimuli and apparatus were the same as those used in Experiment 2.

Procedure. Observers participated in the experiment individually in three sessions conducted on three different days. When observers entered the lab on the first day they were told that they would be participating in an experiment about the relationship between spatial attention and perceptual organization. To that end, they would be asked to keep their eyes fixed in a certain region and to pay attention to another region and report about what they perceived at the attended region. Observers received practice performing these tasks before the experimental trials.

First, they practiced maintaining the after-image within the fixation circle. After-images were induced monocularly in alternate eyes on successive trials so that the after-image used for the first trial decayed while the observer participated in the second trial, etc. After-images were induced foveally. The circular opening in the black opaque shield covering the flash attachment was outlined in flourescent paint to provide viewers with a visible fixation point prior to flash onset. Viewers signaled when they were fixating and the experimenter initiated a 10 ms flash. After the flash, the observers closed their eyes briefly while the experimenter removed the flash apparatus and turned on the strobe light illuminating the paper on which the cubes were drawn. Then, viewing the cube monocularly with the eye in which the after

image had just been generated, observers participated in an average of eight 20 second trials practicing maintaining the after-image within the circle. (Twenty second trials were used in Experiment 3 rather than 30 second trials because the after-images could not be maintained for 30 seconds.)

Next, observers practiced attending to each of the intersections of the cube while maintaining the after-image within the fixation circle. They participated in an average of eight trials, with fixation alternating between the right and left fixation points on each trial.

Next, observers practiced following the opposed-set instructions and pressing one of two keys to indicate which line appeared in front while they maintained the after-image in the fixation circle and their attention on the instructed intersection. On these trials, observers pressed one of two keys to indicate which line appeared forward at any moment and removed their fingers from both keys to report lapses of fixation or attention. There were 16 of these practice trials in which all three tasks were combined, four trials with each hold instruction, one trial per hold instruction at each combination of intersection, fixation point, and eye. Only unbiased cubes were used on these practice trials. Observers were questioned after each practice trial about their ability to perform each of the three tasks when they were combined and they were reminded of the importance of faithful monitoring of the locus of both the after-image and their attention. After these 16 trials, the first session was ended.

The second session began with four more practice trials with the unbiased cubes: two trials with each hold instruction, one per hold

instruction at each intersection. Observers used their right eye for half of the practice trials and their left eye for the other half. They attended to the near intersection for half the practice trials, and to the far intersection for the other half of the practice trials. These practice trials were followed by 16 of the 32 experimental trials.

The third session consisted of four more practice trials followed by the remaining 16 experimental trials. The order of these trials was counterbalanced within subjects.

Results and discussion

Bias influence index. The results are shown in Figure 8. The data are collapsed across fixation point because the previous two experiments showed no effects of fixation point and across right and left eye because we neither expected nor saw any differences in performance with the two eyes. The bias present at the biased intersection was effective: For all subjects on the attend to the biased intersection trials, B was less than zero on hold inconsistent trials and greater than zero on hold consistent trials. Responses to the biased intersection were not obligatory when observers attended to the unbiased intersection, however: In no case was B less than zero on hold inconsistent trials, and in two cases, B was less than zero on hold consistent trials at the unbiased intersection. (The performance of the observers in Experiment 3 on trials during which they attended to the unbiased intersection fell within the range of the performance of the observers in Experiment 2.)

Insert Figure 8 about here

Lapses are shown in Figure 9. The lapses reported by the observers in Experiment 3 fell within the range of the lapses reported by the observers in Experiment 2. For two of the observers $L_e - L_c$ was greater than zero on hold inconsistent trials coupled with attention to the biased intersection, as was found for the observers in Experiment 2. For the third observer (EH), $L_e - L_c$ was approximately zero in this condition. For all observers $L_e - L_c$ was only slightly greater than zero on hold inconsistent trials coupled with attention to the unbiased intersection. The differences were not large, nor were they consistent: For one observer (EH), $L_e - L_c$ was greater on hold consistent trials than on hold inconsistent trials while she attended to the unbiased intersection. She reported later that it was more difficult to keep attention away from the biased intersection on hold consistent trials than on hold inconsistent trials, because she knew she would be more successful at following the hold consistent instructions if she could attend to the biased intersection.

Insert Figure 9 about here

General Discussion

The three experiments reported here show that spatial attention can be directed to a part of an object. The biased intersection did not exert an obligatory influence on the organization perceived at the unbiased intersection, as would be predicted by the object-wide-attention hypothesis. Moreover, we have shown that the allocation of spatial attention within an object affects the perceived organization of that object.

Our results should not be taken as supporting the distinctive features hypothesis of reversal, however. To retain the distinctive features hypothesis in the presence of the evidence presented in these experiments, one would have to suppose that different subparts of the unbiased region favor different interpretations for the cube and that viewers instructed to attend to the unbiased intersection followed the different hold instructions by attending to those different subparts. Clearly, in light of our evidence, the distinctive features hypothesis is forced to predict smaller and smaller sized distinctive features, thereby becoming increasingly ad hoc and increasingly difficult to support experimentally.

Moreover, we do not accept the constructivist assumptions of the distinctive features hypothesis of reversal, namely that the potential interpretations for each of the parts of an object are ordered, and that the parts of the figure are analyzed serially. Recent theory and research in shape recognition suggests that the number of components used to represent shapes is finite (e.g., see Biederman, 1985; 1987). If so, then each representational component occurs in thousands of shape representations and must, therefore,

support large numbers of interpretations. Yet, shape recognition occurs too rapidly to be accounted for by a perceptual mechanism that cycles through large numbers of interpretations for a single part. It is more likely that a stimulus shape is first expressed as a structural description (i.e., a description of the shape's components in their relative locations) and is then matched in parallel against multiple memory representations (cf. McClelland & Rumelhart, 1981; 1986). In what follows, we present a model of how the allocation of attention within an object might influence this process.

<u>The functional equivalence model of spatial attention in object perception</u>.

We propose that when attention is directed to a part of an object, the processing of the stimulus information in the location to which spatial attention is directed is facilitated and the processing of the stimulus information in unattended locations is inhibited. The facilitation takes the form of speeding the processing of the attended region, and hence, of increasing the amount of stimulus information or detail processed in a given time period. The inhibition takes the form of slowing and attenuating (cf. Treisman, 1960) the processing of information in the unattended region. A perceptual interpretation is fitted to an object on the basis of the information apprehended within a given interval of time.

Facilitation of the processing of attended parts of a shape and inhibition of the processing of unattended parts results in an interpretation for the object that depends more upon the attended region than upon unattended regions. We describe the way this occurs using our experimental cubes as examples. When attention is directed to the biased intersection in a partially

biased cube, the processing of the information there is facilitated, thereby speeding the processing of the occlusion and shading details. With this depth information strongly present in the structural description, the structural description of the partially biased cube may be functionally equivalent to that of a more densely biased cube, in that the best matching object representation will be the same as that accessed by a more densely biased cube. On the other hand, when attention is directed to the unbiased intersection, the combined facilitation of the processing of the unbiased intersection and inhibition of the processing of the biased intersection results in a structural description that is functionally equivalent to that of an unbiased cube, in that the set of potentially matching object representations is equivalent to the set accessed by a fully ambiguous cube.

The idea that the allocation of attention in space results in facilitation is not new (cf Erikson & Hoffman, 1972a& b; Posner et al, 1980), although facilitation has not always been construed as an active process. Nor is the proposal that attentional effects include inhibitory processes as well as facilitative processes new (see Maylor & Hockey, 1985; Neill & Westberry, 1987; Posner & Cohen, 1984). The type of inhibition we propose is different from the inhibition proposed by these investigators, however. It is neither an inhibition of competing responses nor an inhibition specific to a previously attended location. Instead, the facilitative and inhibitory operations we propose are similar to those proposed recently by LaBerge and Brown (1989). In their theory, the allocation of attention entails both active facilitation of attended regions and active inhibition of unattended regions. Facilitation

and inhibition are both effected through a filter that operates on the location information in the display, regardless of the content of that location. We assume that spatial attention operates through similar location-directed, content-free mechanisms.

Our theory does not overlap completely with LaBerge and Brown's theory, however, for two reasons. First, we assume that parts of objects can be selected by spatial attention, whereas they assume that entire objects are chosen for facilitation or inhibition. Hence, they cannot account for effects of the allocation of attention to parts of objects such as we have reported here. Second, we propose that the allocation of attention within an object can influence the perceived organization of that object, whereas LaBerge and Brown do not discuss a role for spatial attention in perceptual organization.

"Objects" of spatial attention, organization, and intention.

Spatial attention. The object-wide attention hypothesis concerns the allocation of spatial attention within an object. We have shown that when spatial attention is directed away from regions of the cube that when fixated or attended clearly specify one depth organization, those regions can lose their effectiveness in determining perceived organization. Thus, responses to all parts or properties of the attended object are clearly not obligatory, at least when the task examined is perceptual organization.

In our view, the allocation of spatial attention affects the buildup of information within a structural description of an object. Information from the attended regions builds up in the structural description quickly and information from the unattended regions builds up more slowly. This in turn

affects the process of matching the structural description to an object representation. In Kahneman and Treisman's view, the allocation of spatial attention does not affect the buildup of information within a structural description (or "object file", in their terminology), but only whether or not information of any and all kinds is disseminated from the structural description. In Kahneman and Treisman's theory, then, spatial attention operates later than in our theory. (See Kahneman and Treisman's (1984) summary of the early vs. late selectivity debate.)

The question of how small the focus of attention can be within an object remains. The finding that the partially biased cubes were functionally equivalent to unbiased cubes when spatial attention was directed to the unbiased intersection was obtained in conditions in which the biased and unbiased intersections were separated by at least one degree of visual angle. With attention focused on the unbiased intersection, the biased intersection would fall outside a one degree diameter attentional focus (i.e., the diameter proposed by Erikson & Hoffman (1972) among others). Our results show that, for the purposes of perceptual organization, spatial attention can exclude information at a distance of one degree of visual angle from the location of attentional focus within an object that is larger than one degree of visual angle. We have no evidence that a one degree diameter is the limit for tasks such as these, however.

A related question concerns which types of object properties can be rendered ineffective by the inhibitory effects that accompany the allocation of spatial attention within the object. The experiments reported here show

that the depth cues at the biased intersection can be attenuated when spatial attention is allocated elsewhere. Depth assignment is integral to the process of perceptual organization that we examined, so once the biased intersection is attenuated, the perceived organization is free to alternate between the two potential relative depth interpretations. Future research must determine which object properties (implicated in which processes) are susceptible to attenuation effects and which are not. The relationship between retinal eccentricity, object properties, and spatial attention effects must also be examined (Hochberg & Peterson, 1989).

Organization. In proposing their object-wide attention hypothesis, Kahneman and Henik (1982, p. 181) and Kahneman and Treisman (1984, p. 54) stated that considerations of perceptual organization are indispensible to a theory of attention. We agree, but we submit that careful considertaions of what constitutes an object are equally important to theories of perception and attention. Preattentive processing may determine the boundaries of the object expressed in a structural description (Marr, 1982; Treisman, 1986), but more work must be done before we know what the preattentively defined object entails and how organizational processes interact with the preattentively defined object.

It should be noted that none of the variables identified as preattentive (e.g., similarity of color or of orientation) were involved in defining our cubes as objects. Of course, other preattentive variables may yet be identified; for example, primitive representational components may be identified preattentively (Biederman, 1987). It may turn out that spatial

attention cannot be restricted to a subpart of one of these representational components. In that case, the object-wide attention hypothesis would apply to one level of the hierarchy of objects -- that is, the the level containing "objects" built from one component only. But the characteristics of such one component "objects" and their relationship to objects constructed from a number of structurally relevant components remain to be explored.

The issue of wholes versus parts is important in another regard as well. Observers in our experiments and in previous opposed-set experiments tended to report that the perceived organization of the cube reversed as a whole. That is, the depth relationships at the biased intersection did not remain fixed when the unbiased intersection reversed, as might be implied by the term "piecemeal reversal." These phenomenological reports suggest that, even though spatial attention need not be allocated object-wide, perceptual organization may be fit object-wide (Koffka, 1935). But this is a tentative statement and must remain so until experiments directed specifically to that question are conducted. Whether piecemeal or holistic reversals are obtained may depend upon the number and type of components from which the object is constructed, upon the minimum grain size of perceptual organization (Hochberg, 1982), and upon the type of processes involved in the perceptual organization of the object. For example, some minimum number of components may be necessary before any interpretation can be seen reliably (Biederman, 1987), so piecemeal reversals may be possible only when an object can be decomposed into sub-objects each composed of a minimum number of components.

Intention. When spatial attention location was held constant at the unbiased intersection, viewers were able to successfully follow the instructions to hold both alternatives. Thus, these experiments suggest that perceptual intentions need not operate through the allocation of attention in space. Although the results obtained when attention was directed to the biased intersection suggest that the allocation of attention in space may be an effective avenue through which intentions might operate in other situations, the results obtained when attention was directed to the unbiased intersection suggest that intention can operate through other mechanisms as well.

One possible mechanism through which intentions might operate is the top-down activation of canonical object representations (Peterson et al, 1989). A question raised by the current set of experiments is whether perceptual intentions operate effectively through the top-down activation of holistic object representations only, and not part representations. (Remember that the perceived organization of the cube reversed holistically.) If so, that might imply that response selectivity -- in this case, the fitting of a perceptual organization -- must be object-wide, even though input selectivity -- the focus of spatial attention -- need not be. Further research is needed before any such conclusions can be reached, but the present set of experiments clearly show that spatial attention and intention are separable modes of selectivity.

References

Alpern, M. (1972). Effector mechanisms in vision. In J.W. Kling & L.A. Riggs (Eds.) Experimental Psychology. New York: Holt, Rinehart, and Winston.

Beck, J. (1972). Similarity grouping and peripheral discriminability under uncertainty. American Journal of Psychology, 85, 1-19.

Beck, J. (1982). Textural segmentation. In J. Beck (Ed.), Organization and representation in perception. LEA: Hillsdale

Biederman, I. (1985). Human image understanding: Recent experiments and a theory. Computer Vision, Graphics, and Image Processing, 32, 29-73.

Biederman, I. (1987). Recognition by components: A theory of human image understanding. Psychological Review, 94, 115-147.

Chastain, G. & Burnham, C.A. (1975). The first glimpse determines the perception of an ambiguous figure. Perception and Psychophysics, 17, 221-224.

Downing, C.J. (1988). Epectancy and visual-spatial attention: Effects on perceptual quality. Journal of Experimental Psychology: Human Perception and Performance, 14, 188-202.

Downing, C.J. & Pinker, S. (1985). The spatial structure of visual attention. In M.I. Posner & O.S.M. Marin (Eds.), Attention and performance, XI. Hillsdale: LEA.

Ellis, S.R. & Stark, L. (1979). Eye movements during viewing of Necker cubes. Perception, 7, 575-581.

Epstein, W. (1982). Percept-percept couplings. Perception, 11, 75-83.

Erikson, C.W. & Hoffman, J.E. (1972a). Temporal and spatial characteristics of selective encoding from visual displays. Perception and Psychophysics, 12, 201-204.

Erikson, C.W. & Hoffman, J.E. (1972b). Some characterisitics of selective attention in visual perception determined by vocal reaction time. Perception and Psychophysics, 11, 169-171.

Gogel, W.C. (1980). The sensing of retinal motion. Perception and Psychophysics, 28, 155-163.

Gogel, W.C. & Tietz, J. (1974). The effect of perceived distance on perceived movement. Perception and Psychophysics, 16, 70-78.

Hochberg, J. (1968). In the mind's eye. In R.N. Haber (Ed.), Contemporary theory and research in visual perception. New York: Appleton-Century-Crofts.

Hochberg, J. (1974). Higher-order stimuli and interresponse coupling in the perception of the visual world. In R.B MacLeod & H.L. Pick (Eds.), Perception: Essays in honor of James J. Gibson. Ithaca: Cornell University Press.

Hochberg, J. (1982). How big is a stimulus? In J. Beck (Ed.), Organization and representation in perception. Hillsdale: LEA.

Hochberg, J. & Hay, J (1956). figural after-effect, after-image, and physiological nystagmus. American Journal of Psychology, 69, 480-482.

Hochberg, J. & Peterson, M.A. (1987). Piecemeal Perception and cognitive components in object perception: Perceptually coupled responses to moving objects. Journal of Experimental Psychology: General, 116, 370-380.

Hochberg, J. & Peterson, M.A. (1989). Bildanalyse in der menschlichen informationverarbeitung (Pictures in the mind's eye: Images in our perception of world and art). In M. Schuster & B.P. Woschek (Eds.), Nonverbale kommunikation durch bilder. Stuttgart: Verlag fur Angewandte Psychologie.

Hughes, H.C. & Zimba, L.D. (1985). Spatial maps of directed visual attention. Journal of Experimental Psychology: Human Perception and Performance, 11, 409-430.

Kahneman, D. & Henik, A. (1977). Effects of visual grouping on immediate recall and selective attention. In S. Dornic (Ed.), Attention and performance, VI. Hillsdale: LEA.

Kahneman, D. & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J.R. Pomerantz (Eds.), Perceptual organization. Hillsdale: LEA.

Kahneman, D. & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D.R. Davies (Eds.), Varieties of attention. New York: Academic Press.

Koffka, K. (1935). Principles of gestalt psychology. New York: Harcourt, Brace.

LaBerge, D. & Brown, V. (1989). Theory of attentional operations in shape identification. Psychological Review, 96, 101-124.

Maylor, E.A. & Hockey, R. (1985). Inhibitory component of externally controlled covert orienting in space. Journal of Experimental Psychology: Human Perception and Performance, 11, 777-787.

Marr, D. (1982). Vision. San Francisco: W.H. Freeman.

McClelland, J.L. & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. Psychological Review, 88, 375-407.

McClelland, J.L. & Rumelhart, D.E. (1986). Parallel distributed processing. Vol 1: Foundations. Cambridge: MIT.

Merikle, P.M. (1980). Selection from visual persistence by perceptual groups and category membership. Journal of Experimental Psychology: General, 109, 279-295.

Merikle, P.M. & Gorewich, N.J. (1982). Spatial selectivity in vision: Field size depends upon noise size. Bulletin of the Psychonomic Society, 14, 343-346.

Neill, W.T. & Westberry, R.L. (1987). Selective attention and the suppression of cognitive noise. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 327-334.

Neisser, U. (1967). Cognitive psychology. New York: Appleton-Century-Crofts.

Olson, R. & Ohrbach, J. (1966). Reversibility of the Necker cube: VII. parts of the picture contributing to the perception of reversals. Perceptual and Motor Skills, 22, 623-629.

Peterson, M.A. (1984). Measures of selective components in perceptual organization. Unpublished doctoral dissertation, Columbia University, New York.

Peterson, M.A. (1986). Illusory concomitant motion in ambiguous stereograms: Evidence for nonstimulus contributions to perceptual organization. Journal of Experimental Psychology: Human Perception and Performance, 12, 50-60.

Peterson, M.A., Harvey, E., & Weidenbacher, H.B. (1989). Canonical descriptions, reference frames, and viewer intention in figure-ground perception. Manuscript in preparation.

Peterson, M.A. & Hochberg, J. (1983). Opposed-set measurement procedure: A quantitative analysis of the role of local cues and intention in form perception. Journal of Experimental Psychology: Human Perception and Performance, 9, 183-193.

Peterson, M.A. & Hochberg, J. (1989). Necessary considerations for a theory of form perception: A theoretical and empirical reply to Boselie and Leeuwenberg (1986). Perception, 18, 105-119.

Phieffer, C.H., Eure, S.B., & Hamilton, C.B. (1956). Reversible figures and eye movements. American Journal of Psychology, 69, 452-455.

Pomerantz, J.R. (1983). Global and local precedence: Selective attention in form and motion perception. Journal of Experimental Psychology: General, 112, 516-540.

Posner, M.I. & Cohen, Y. (1984). Components of visual orienting. In H. Bouna & D.G. Bouwhis (Eds.), Attention and performance, X. Hillsdale: LEA.

Posner, M.I., Snyder, C.R.R., & Davidson, B.J. (1980). Attention and the detection of signals. Journal of Experimental Psychology: General, 109, 160-174.

Post, R.B. & Leibowitz, H.W. (1982). The effect of convergence on the vestibulo-ocular reflex and implications for perceived movement. Vision Research, 22, 461-465.

Price, J.R. (1967). Perspective duration of a plane reversible figure. Psychonomic Science, 9, 623-624.

Pritchard, R.M. (1958). Visual illusions viewed as stabilized retinal images. Quarterly Journal of Experimental Psychology, 10, 77-81.

Remington, R.W. (1980). Attention and saccadic eye movements. Journal of Experimental Psychology: Human Perception and Performance, 6, 726-744.

Sperling, G. (1960). The information available in brief visual presentations. Psychological Monograph: General and Applied, 74, 1-29. (Whole No. 498).

Treisman, A. (1960). Contextual cues in selective listening. Quarterly Journal of Experimental Psychology, 12, 242-248.

Treisman, A. (1969). Strategies and models of selective attention. Psychological Review, 76, 282-299.

Treisman, A. (1986). Properties, parts and objects. In K. Boff, L. Kauffman, & J. Thomas (Eds.), Handbook of perception and human performance. New York: Wiley.

Treisman, A. (1988). Features and objects: The fourteenth Bartlett Memorial Lectures. The Quarterly Journal of Experimental Psychology, 40, 201-237.

Treisman, A. & Gelade, G. (1980). A feature integration theory of attention. Cognitive Psychology, 12, 97-136.

Tsal, Y. & Kolbert, L. (1985). Disambiguating ambiguous figures by selective attention. Quarterly Journal of Psychology, 37A, 25-37.

von Wright, J.M. (1968). Selection in immediate visual memory. Quarterly Journal of Experimental Psychology, 20, 62-68.

Figure Captions

Figure 1. One of the partially-biased Necker cubes used by Peterson and Hochberg (1983). At 1, the "biased" intersection, occlusion and shading specify that the horizontal line is in front of the vertical line. At 2, the "unbiased" intersection, no information has been added to specify that either line is in front.

Figure 2. The partially-biased Necker cubes used in Experiment 1. The biased intersection of the cubes in the top row specifies that the cube is facing downwards and to the left. The biased intersection of the cubes in the bottom row specifies that the cube is facing upwards and to the right. The three fixation points used in Experiment 1 are indicated in the figures.

Figure 3. Results of Experiment 1: Intention index (_I_) averaged across observers as a function of hold instruction (consistent or inconsistent with bias) at each intersection (biased and unbiased).

Figure 4. Lapses in Experiment 1 as a function of hold instruction (consistent and inconsistent with bias) and intersection (biased and unbiased).

Figure 5. The cubes used in Experiment 2. A and B the experimental cubes, which are biased toward the two orientations at the top intersection. C is the control cube, which remains unbiased at both intersections. All cubes were presented with each of the three fixation points used in Experiment 1.

Figure 6. The bias influence index (_B_) averaged across observers as a function of hold instruction and intersection in Experiment 2.

Figure 7. $(L_e - L_c)$ averaged across observers and trials for each fixation point, intersection, and hold instruction in Experiment 2.

Figure 8. Bias influence index (*B*) for each of three observers (averaged across fixation point and eye) as a function of hold instruction and intersection in Experiment 3.

Figure 9. $(L_e - L_c)$ as a function of hold instruction and intersection for each of the three subjects in Experiment 3.

Author Notes

This research was supported in part by a grant to the first author from The National Science Foundation (BNS-8810997) and the Air Force Office of Scientific Research.

Footnotes

[1]The biased and unbiased intersections were separated by less than 2 degrees of visual angle, so in any case, the previous results provide evidence for the independent organization of parts of an object that is larger than two degrees of visual angle (Peterson & Hochberg, 1988).

[2]All effects, where applicable, are replicated with total durations.

FIGURE 1

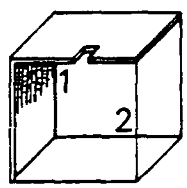


FIGURE 2

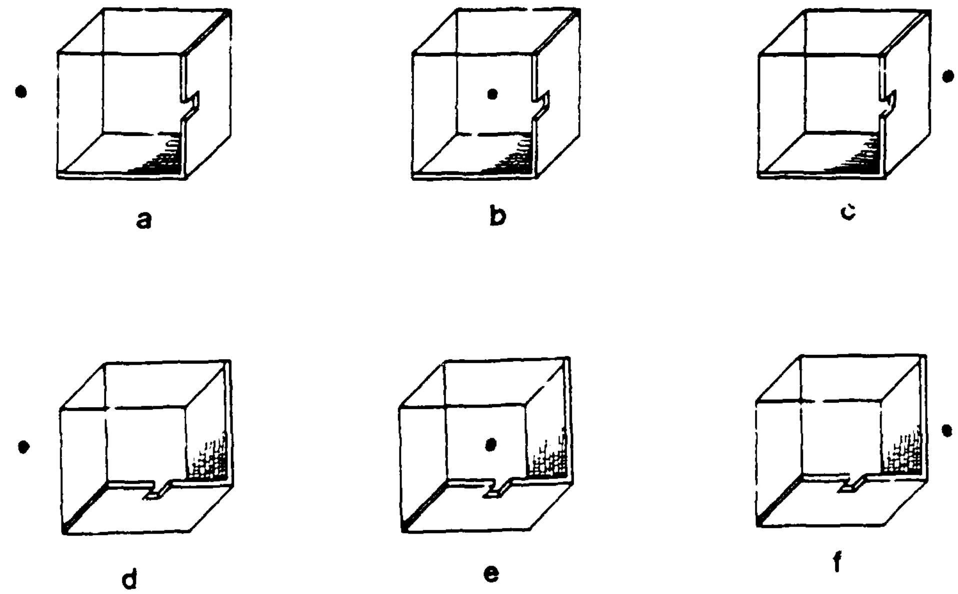


FIGURE 3

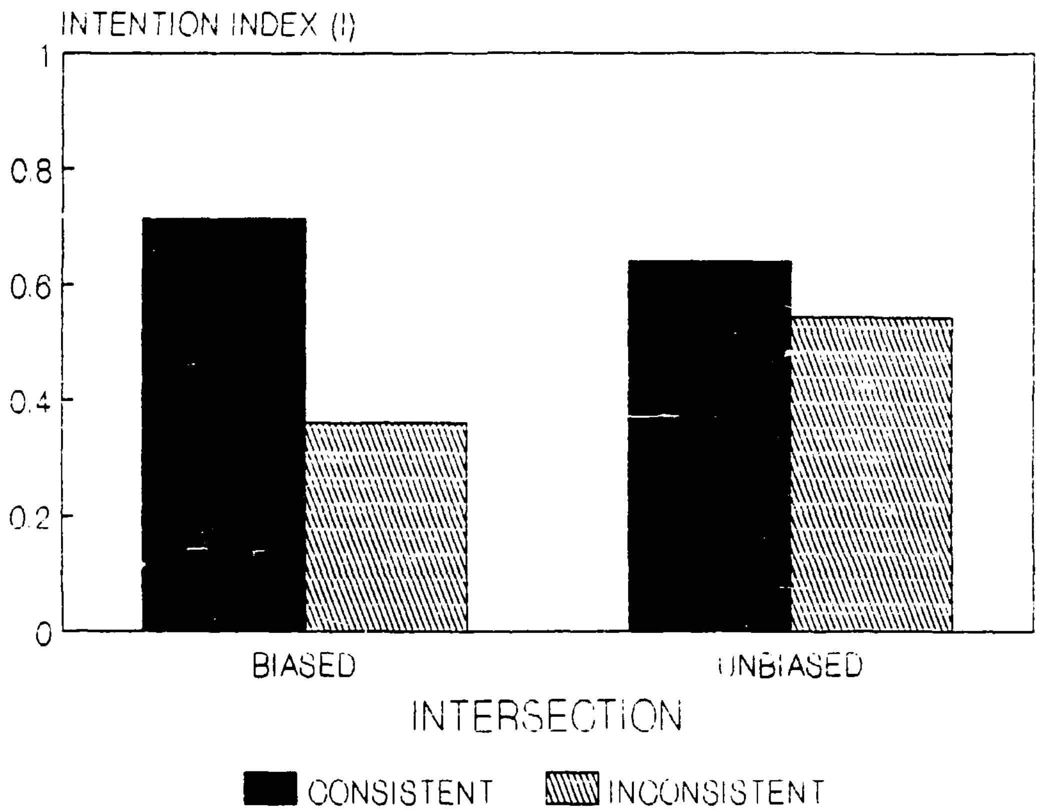

INTENTION INDEX (I)
1
0.8
0.6
0.4
0.2
0
BIASED
UNBIASED
INTERSECTION
CONSISTENT
INCONSISTENT

FIGURE 4

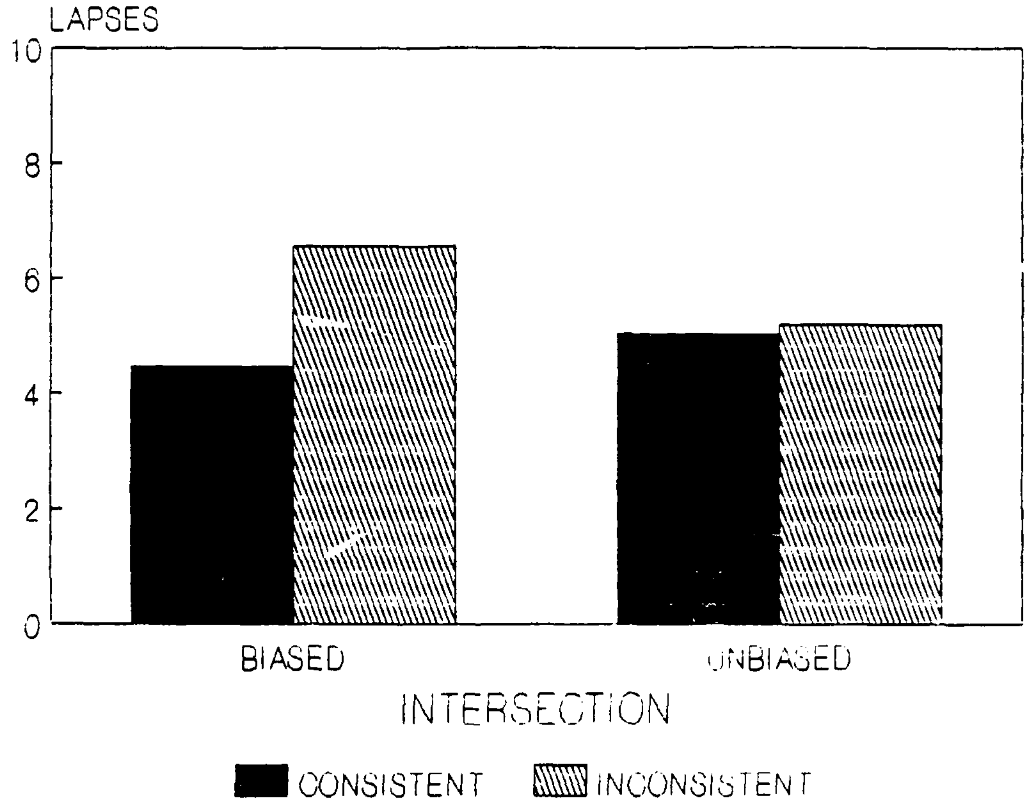
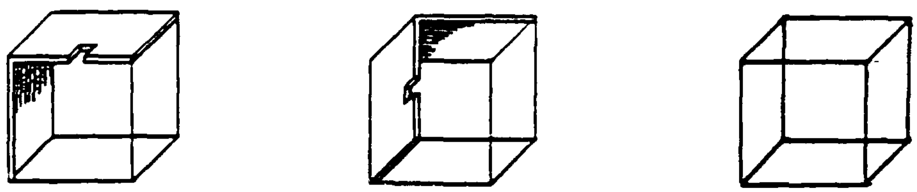

LAPSES
10
8
6
4
2
0
BIASED
UNBIASED
INTERSECTION
CONSISTENT
INCONSISTENT

FIGURE 5



a b c

FIGURE 6

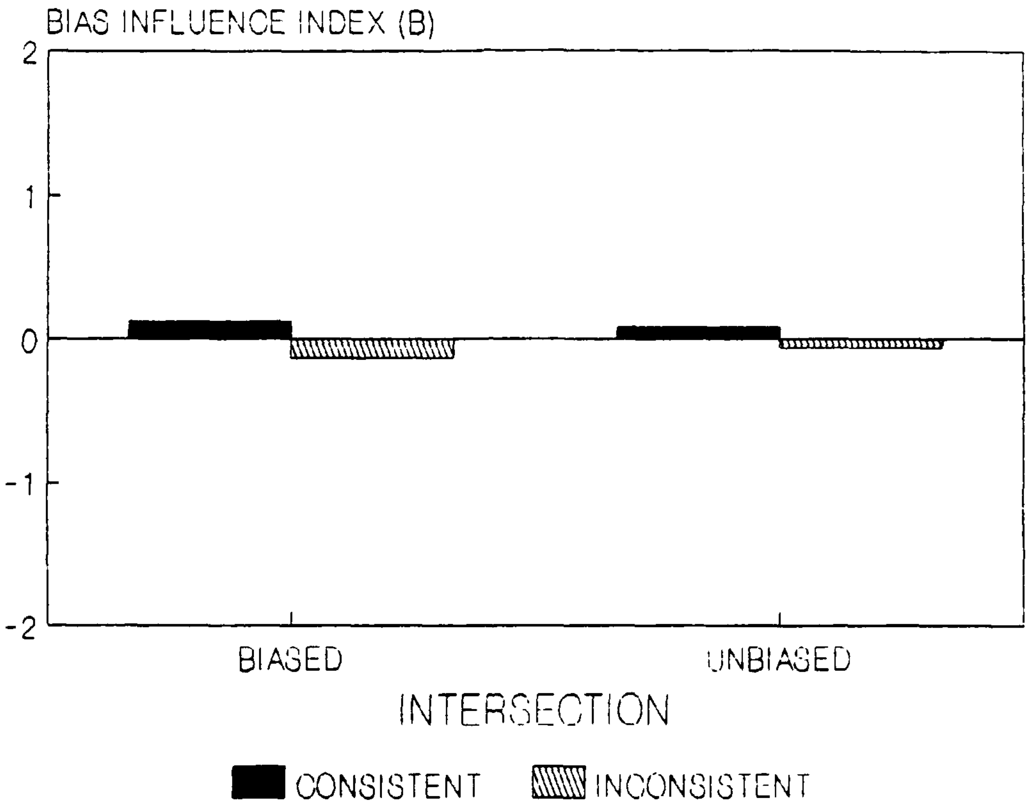
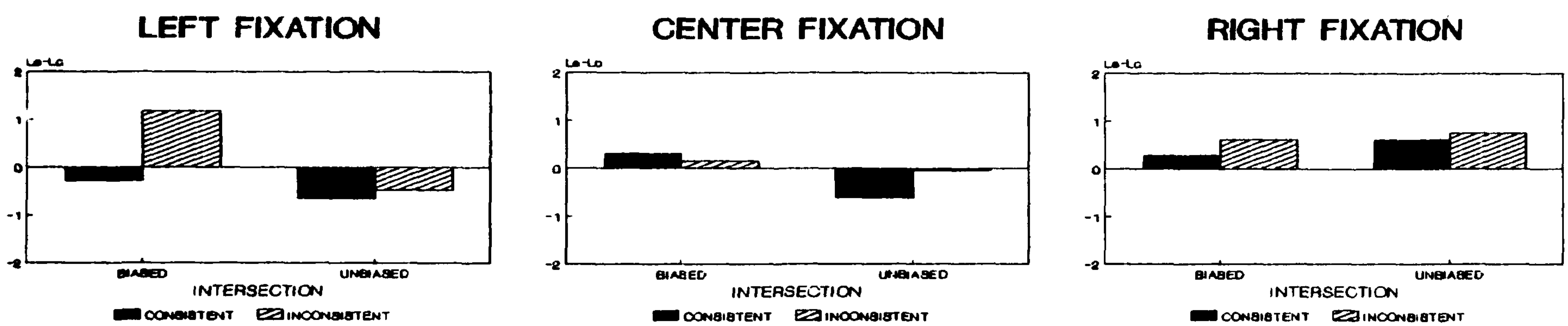

BIAS INFLUENCE INDEX (B)
2
1
0
-1
-2
BIASED
UNBIASED
INTERSECTION
CONSISTENT
INCONSISTENT

FIGURE 7



LEFT FIXATION
Ls-Lc
2
1
0
-1
-2
BIASED
UNBIASED
INTERSECTION
CONSISTENT
INCONSISTENT
CENTER FIXATION
Ls-Lc
2
1
0
-1
-2
BIASED
UNBIASED
INTERSECTION
CONSISTENT
INCONSISTENT
RIGHT FIXATION
Ls-Lc
2
1
0
-1
-2
BIASED
UNBIASED
INTERSECTION
CONSISTENT
INCONSISTENT

FIGURE 8

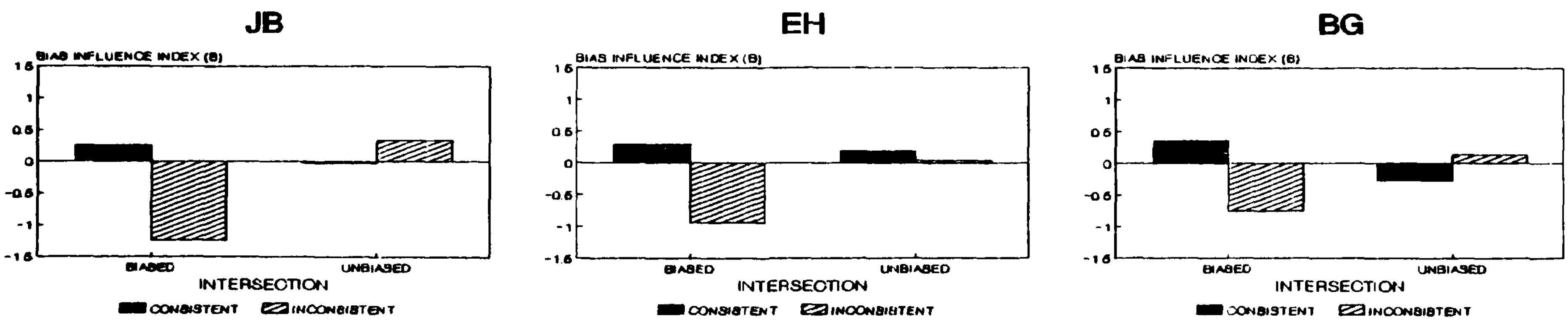
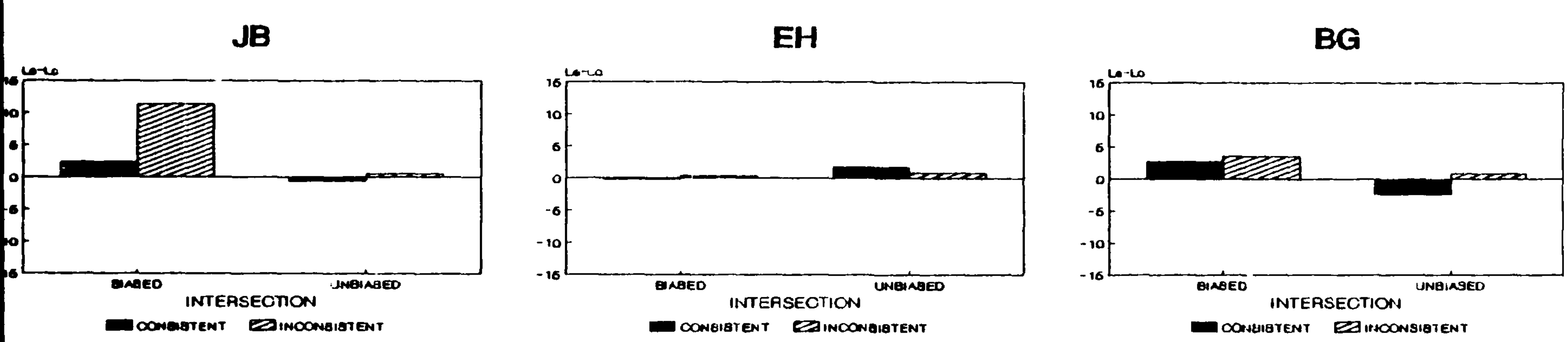

JB
EH
BG
BIAS INFLUENCE INDEX (B)
1.5
1
0.5
0
-0.5
-1
-1.5
BIASED
UNBIASED
INTERSECTION
CONSISTENT
INCONSISTENT

FIGURE 9


JB
EH
BG
Le-Lo
15
10
5
0
-5
-10
-15
BIASED
UNBIASED
INTERSECTION
CONSISTENT
INCONSISTENT

in press,
Journal of Mental Imagery

Vividness and Control of Mental Imagery: A Psychometric Analysis

John F. Kihlstrom, Martha L. Glisky,
Mary A. Peterson, Erin M. Harvey, and Patricia M. Rose
University of Arizona

Running Head: Imagery Vividness and Control

Address for Correspondence: John F. Kihlstrom
Amnesia and Cognition Unit
Department of Psychology
University of Arizona
Tucson, Arizona 85721

Abstract

In a study of the relations between vividness and control of mental imagery, a total of 2083 subjects completed the shortened form of Betts' QMI and Gordon's TVIC; another 730 subjects completed Marks' VVIQ and the TVIC. Distributions of responses on all three scales were highly skewed, with most subjects reporting at least moderately clear and vivid images. Factor analysis yielded 7 factors in QMI, 4 factors in VVIQ, and 2 to 4 factors (depending on the sample) in the TVIC. Although VVIQ was intended as an expansion of the visual imagery subscale of QMI (QMI/V), TVIC was significantly more highly correlated with VVIQ than with QMI/V -- a difference that was not due to sampling effects, differential reliability, or restriction of range. The available imagery questionnaires confuse the dimensions of vividness and control, fail to apply coherent definitions of either attribute of imagery, and are relatively insensitive to individual differences in imagery ability.

Vividness and Control of Mental Imagery:
A Psychometric Analysis

Over the last quarter century, the study of mental images has progressed from a reliance on introspective self-reports of imaginal experience to analyses of the processes underlying their manipulation and use on experimental tasks (for historical reviews, see Holt, 1964; Hilgard, 1981; Sheehan, 1978; for accounts of some contemporary issues, see Block, 1981; Kosslyn, 1980; Shepard & Cooper, 1982; Yuille, 1983). In the course of this research, various questionnaire measures have been devised in an attempt to capture individual differences in imagery skill. Among the most popular of these are the shortened form of Betts' Questionnaire upon Mental Imagery (QMI; Betts, 1909; Sheehan, 1967), Marks' Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973), and Gordon's Test of Visual Imagery Control (TVIC; Gordon, 1949).

Analyses of data derived from these questionnaires have yielded a number of interesting facts concerning the phenomenal experience of mental imagery (for recent reviews, see Sheehan, Ashton, & White, 1983; Tower & Singer, 1981; White, Sheehan, & Ashton, 1977). For example, there appears to be a strong general factor of imagery vividness, such that vividness in one domain (such as vision) is positively correlated with vividness in other domains (e.g., gustation). However, detailed analyses also reveal a small number of specific factors corresponding to the physiological substrates of the various sensory modalities, such as chemical (e.g., olfactory and gustatory) and mechanical (e.g., auditory and tactile). Self-ratings of vividness tend to be somewhat skewed, with most subjects claiming to experience at least moderately vivid

images, and vanishingly few subjects reporting no imagery at all. Finally, images in some modalities (e.g., visual, auditory) tend to be more vivid than those in others (e.g., gustatory, olfactory). While this pattern may reflect little more than the relative familiarity of the various modalities, it also parallels the distribution of imagery in other domains. For example, visual and auditory imagery predominates in autobiographical memory, while olfactory and gustatory imagery is relatively absent (e.g. Kihlstrom & Harackiewicz, 1982).

While questionnaire measures seem to reveal clear individual differences in vividness and control of mental imagery (e.g., White, Ashton, & Brown, 1977), it has been difficult to show that these differences have any functional significance. Although Marks (1983) reviewed results showing that VVIQ scores correlated significantly with performance on a variety of tasks involving perception and memory, other reviewers have been less sanguine (e.g , Ernest, 1977; Finke, 1980; Katz, 1983; Richardson, 1980; Sheehan, Ashton, & White, 1983; Tower & Singer, 1981; White, Sheehan, & Ashton, 1977). For example, while scores on the various imagery scales (such as QMI and VVIQ) often correlate highly with each other, establishing a kind of convergent validity, they do not correlate reliably with performance on tests of spatial ability, mental rotation, or image scanning. Reisberg and his colleagues have consistently found that VVIQ scores predict subjects' ability to remember details of the visual appearances of objects; somewhat surprisingly, however, the correlation consistently obtained is <u>negative</u>, meaning that vivid imagers are <u>less</u> accurate than their counterparts with more vague and fragmentary images (for a review, see Reisberg & Heuer, 1989a, 1989b).

In the course of conducting research on the relationship between visual perception and imagery (Peterson, Kihlstrom, Rose, Glisky, & Harvey, 1989), we have administered various scales measuring the vividness and control of mental imagery to large pools of subjects. When it came to selecting subjects for our experiments, we were struck by the absence of contemporary norms for interpreting scale scores. In this paper, we briefly report norms for several of these instruments on a new generation of college students, and point out some heretofore unappreciated aspects of vividness, control, and the relationships between them.

Method

Subjects and Procedure

The data analyzed in this report was taken from three samples of college students enrolled in introductory psychology courses, who received either the QMI or the VVIQ, each coupled with the TVIC, during a routine survey session conducted near the beginning of the semester. A total of 1439 subjects at the University of Wisconsin completed the QMI and TVIC in the spring semester of 1986; another 584 subjects completed these same questionnaires at the University of Arizona in the fall semester of 1987. Finally, a total of 730 University of Arizona students completed the VVIQ and TVIC in the spring semester of 1989. In the Wisconsin sample, the subjects received extra credit through the research participation option of the course. In all samples, the subjects were informed that their responses to the questionnaires might make them eligible to participate in experiments offered under the research participation requirement of the course. The subjects completed the

questionnaires at their own pace, usually requiring less than 10 minutes for the task.

Materials

The shortened form of Betts' (1909) QMI (Sheehan, 1967) consists of a total of 35 items, with five items in each of seven subscales tapping visual, auditory, tactile (cutaneous), kinesthetic, gustatory, olfactory, and organic (whole body) imagery. For example, subjects are asked to imagine "the sun as it is sinking below the horizon", and then to rate the vividness of the resulting image on 7-point scale ranging from "Perfectly clear and as vivid as the actual experience" to "No image present at all, only 'knowing' that you are thinking of the object". In this research, we were particularly interested in the visual imagery subscale of the QMI, henceforth referred to as QMI/V.

Marks' (1973) VVIQ was intended to be an expansion of the QMI/V. It consists of 16 items, 4 items relating to each of 4 different images (viz., that of a relative or friend, a rising sun, a familiar shop, a country scene), and employs a 5-point scale with the same endpoints as the QMI.

Gordon's (1949) TVIC is concerned with subjects' ability to voluntarily manipulate or control mental images, irrespective of their vividness. It consists of a series of 12 items, all involving an automobile, which goes through various transformations (e.g., stopped, lying upside down, climbing up a hill, falling off a bridge). For each item, subjects are asked to indicate whether they can imagine the requested scene. The subjects at Wisconsin and Arizona who completed the TVIC along with the QMI made their responses on Gordon's original 3-point scale labelled "Yes", "Unsure", and "No". For the

Arizona subjects who completed the TVIC with the VVIQ, however, the scale was expanded to 5 points, retaining the original anchor labels, in an attempt to facilitate the expression of individual differences.

Results

Distribution of Imagery Vividness and Control

QMI. Of the 2083 subjects who received the QMI in both institutions, a total of 2036 completed the scale. Table 1 shows the means and standard deviations, on each of the seven subscales of the questionnaire. Since low scores on the test reflect vivid mental imagery, and the midpoint on each subscale corresponds to a score of 20, it is apparent that the subscale distributions are substantially skewed, with the average subject claiming to experience at least moderately clear and vivid images (i.e., a value of 3 or more on the 7-point scale) in each domain. Depending on the subscale, only 1%-2% of the subjects reported that their images were vague and dim (i.e., a value of 2 or less on the 7-point scale) at best.

A within-subjects analysis of variance yielded a significant main effect of subscale, $F(6, 12{,}210) = 212.67$). "Organic" imagery, involving such experiences as fatigue and drowsiness, was the most vivid. Of the individual imagery modalities, vision yielded the most vivid images, olfaction the least.

<<<<<Place Table 1 About Here>>>>>

VVIQ. A total of 730 subjects returned completed questionnaires. The mean score was 33.31 ($SD = 11.76$), compared to a midpoint score corresponding to 48. Again, the skewness of the distribution indicates that most subjects claimed that their visual images were at least moderately clear and vivid

(i.e., a value of 3 or more on the 5-point scale). Fewer than 3% of the subjects reported that their images were vague and dim, or that they had no imagery at all.

TVIC. A total of 2075 subjects completed the 3-point version, and another 730 subjects completed the 5-point version of the questionnaire. The mean scores were 16.73 (SD = 4.68) against a midpoint of 24, and 23.61 (SD = 8.61) against a midpoint of 36, respectively. Again, most subjects reported that they could experience the images requested. Fewer than 7% of the subjects, overall, reported that they could not produce the images as requested, or doubted that they could do so.

Correlations among Imagery Measures

QMI and TVIC. Table 2 presents the correlations among the QMI subscales, and between each subscale and the 3-point TVIC, based on the 2029 subjects who had complete data on both questionnaires. All the correlations were positive. However, the correlations between TVIC and the subscales of the QMI were all relatively low; the correlation between the QMI/V and TVIC was only $r = 0.25$.

<<<<<Place Table 2 About Here>>>>>

A principal-components factor analysis of the individual QMI and TVIC items, with orthogonal rotation, yielded 11 factors accounting for 57% of the total variance. The first seven factors corresponded closely to the seven subscales of the QMI, confirming the findings of White, Ashton, & Law (1974). Table 3 shows the remaining four factors, which were comprised of TVIC items: Factors 10 and 11 referred to the automobile at rest in a normal position, in color (Factor 10) or not (Factor 11); Factor 8 referred to the car in normal motion; Factor 9 to the car in unusual positions or motions. Ashton and White

(1974) also found that the TVIC was factorially complex, with three factors corresponding to whether the car was standing still, in motion, or in a bizarre position (whether standing or moving).

<<<<<Place Table 3 About Here>>>>>

VVIQ and TVIC. The correlation between VVIQ and the 5-point TVIC was $r = 0.45$. This value was significantly higher than the corresponding correlation between the TVIC and the QMI/V ($r = 0.25$, $p < 0.001$). In interpreting this difference, it should be noted that most of the QMI data was collected at Wisconsin, whereas all the VVIQ data was collected at Arizona, thus raising the possibility of subject selection effects. However, the correlation between TVIC and the QMI/V, when calculated on the Arizona subsample ($N = 584$) alone, was $r = 0.21$. Thus, cohort differences cannot account for the differences observed between the correlations.

In addition, it should be noted that the QMI/V consists of only 5 items, while the VVIQ consists of 12 items (in fact, the first cluster of four VVIQ items is identical to the first four QMI/V items, while the first item in the second cluster of four VVIQ items is identical to the fifth QMI/V item). According to psychometric theory, the correlation between two variables increases with the reliability of their measurement, and as a rule longer scales are more reliable than shorter ones. Accordingly, the correlation was recomputed between TVIC and a VVIQ subscale consisting only of the first five items of the questionnaire (i.e., the items corresponding to the QMI/V subscale). This correlation was $r = 0.36$, which remains significantly higher than the $r = 0.21$ obtained between QMI/V and TVIC in the Arizona subsample $p <$

0.05). Thus, the difference between the correlations cannot be accounted for by differences between QMI and VVIQ in reliability or restriction of range.

On the other hand, the subjects in the QMI/TVIC sample completed the original three-point version of TVIC, while those in the VVIQ/TVIC sample completed a five-point revision. Since most people report that they are able to produce the requested images on TVIC, it may be that the three-point version restricted the range of possible scores available to subjects in the QMI/TVIC sample. In order to investigate this possibilities, the five-point scales for the TVIC items in the VVIQ/TVIC sample were collapsed to three-point ones by combining the two positive scale values (i.e., 1 and 2), and the two negative scale values, (i.e., 4 and 5) together. The resulting distribution of TVIC scores closely resembled that obtained in the QMI/TVIC sample. Nevertheless, the correlation fell only to $\underline{r} = 0.41$ with the full VVIQ score, and to $\underline{r} = 0.32$ with the VVIQ subscale corresponding to QMI/V. Both values are significantly higher than the value of $\underline{r} = 0.21$ obtained between TVIC and QMI/V in the Arizona subsample ($\underline{p} < 0.05$). Thus, the different correlations obtained between TVIC and QMI/V, and between TVIC and VVIQ, cannot be accounted for by restrictions in the range of TVIC scores.

A principal-components factor analysis of the individual VVIQ and TVIC items, with orthogonal rotation, yielded six factors accounting for 64% of the total variance. Four of the factors (Factors 2, 3, 5, and 6) corresponded to the four content clusters of the VVIQ. Table 3 shows that the remaining two factors were comprised of TVIC items -- Factor 1 consisting of items referring to the automobile in normal positions or motions, Factor 4 to the car in unusual positions or motions.

Discussion

One of the most striking findings to emerge from our analyses is that the distribution of mental imagery ability, whether construed in terms of questionnaire-based estimates of vividness or control, is highly skewed. While scores on these scales might be inflated somewhat by the desire of subjects in an experimental pool to participate in further experiments, distributions of similar form have been obtained under other conditions of testing. Therefore, as Sheehan (1967) has noted, it appears that few subjects lack the ability to produce the images requested by these questionnaires. If a self-rating of "moderately clear and vivid" is taken as the threshold for affirming that one has formed a mental image, then the endorsement rates for the QMI items vary from 48.4% to 93.8%, with a mean of 80.82% (SD = 10.78); and for VVIQ items from 76.8% to 94.4% (M = 87.26, SD = 4.78)[1]. Fewer than 10% of the subjects (far fewer on QMI and VVIQ) yielded scores indicating that their images were vague and dim, or that they had no imagery at all. Psychometric principles of test construction usually require a distribution of item difficulties, averaging about 50%. Compared against this standard, the QMI and VVIQ are very "easy" tests.

This psychometric fact has serious implications for the use of these tests in selecting subjects for experiments on mental imagery. In the first place, the common practice in imagery experiments of administering the tests to a small random sample of subjects, and then correlating imagery test scores with performance on some experimental task, will rarely reveal an effect of individual differences. This is because the variance of imagery scores in any unselected subject group is likely to be very low. And if the variance of

imagery scores is low, then such scores cannot serve as strong predictors of performance on some experimental test.

The obvious solution to this problem is to select experimental subjects from the extremes of the distribution of the test scores. In our own research, for example, we compared subjects scoring in the first and fourth quartiles on both VVIQ and TVIC (Peterson et al., 1989). Unfortunately, the skewed distribution of imagery scores vitiates even this solution. For example, in our data the cutpoint for the VVIQ "low vividness" group was 39 on a 60-point scale. If a subject reported experiencing a "moderately clear and vivid" image in response to each item, he or she would have a score of 36, while a rating of "vague and dim" on every item would produce a score of 48. Thus, even with an extremely large population of subjects to choose from, the group of ostensibly poor imagers actually included many subjects whose ability to form at least some images may have been fairly good. The typical failure of individual differences on imagery vividness to reliably predict performance on experimental tests of imagery may be due to the failure of the available tests to capture the full range of imagery ability. Alternatively, it may be that even "vague and dim" images are sufficient for adequate performance on most experimental imagery tasks.

Another striking feature of our results was the differential correlation of the TVIC, ostensibly a measure of imagery control, with the VVIQ and the QMI/V. The VVIQ has generally been considered to be a more reliable extension of the QMI/V, but in our data, the TVIC correlated much more highly with VVIQ scores than with QMI/V -- a difference that cannot be accounted for by cohort effects, differential reliability, or restrictions in range. It appears that the VVIQ, while intended as an extension of QMI/V and thus as a measure of

imagery vividness, may also tap some aspects of controllability. For example, the four successive items in each of the four clusters of the VVIQ may be interpreted as manipulations or transformations of a single basic image of a relative or friend, a sunrise, a shop, and a landscape. Although the subjects are asked to rate the vividness of each image, most VVIQ items represent the manipulation of a prior image, much in the manner of the TVIC items.

In retrospect, it appears that the QMI, VVIQ, and TVIC do not provide unambiguous assessments of either the vividness and controllability of mental images. As noted previously, in many items of the VVIQ the subject is transforming the image in a way that seems very similar to the control measured by the TVIC. Likewise, many of the TVIC questions, such as those referring to the car in a stationary position, seem to pertain to vividness at least as much as they do to control. Even assuming that the questionnaire items measure what they are intended to measure, important questions remain unaddressed. Within the vividness dimension, for example, the complexity of the image requested, and the precision of the attendant details required, have not been controlled. For example, the items in some clusters of VVIQ refer to very complex images, while others refer to more schematic, prototypical scenes. And it is unclear whether vividness should be construed in terms of color brightness, clarity of contours, preservation of precise detail, or some other feature of the image. Within the controllability dimension, many of the TVIC items can be endorsed by calling up a new image, rather than by manipulating one that has been established previously (see also Chambers & Reisberg, 1985). And various aspects of manipulation, for example taking a different perspective on a scene, focusing on some precise detail, or transforming one or more features, are not clearly distinguished.

Even so, considerations of psychometric theory cast doubt on <u>any</u> relationship between vividness and controllability measured with the currently available scales. Arguably, vividness and controllability are conceptually quite different: one could have a very vivid mental image but be unable to manipulate it, for example. And, indeed, the correlations obtained between measures of vividness and control are relatively low, even if they are positive and statistically significant.[2] Unfortunately, the peculiarities of the distributions of QMI/V, VVIQ, and TVIC scores also call this empirically observed independence into question. Correlations express the degree to which variability in one measure can be accounted for by variability in another, and the correlation between two test scores is constrained by the variability of the scores themselves. In the present instance, the variability of QMI/V, VVIQ, and TVIC scores is relatively low: the vast majority of subjects score high on the dimensions measured. The true correlation between vividness and controllability can only be determined with measures that include items tapping a wider range of difficulty levels.

In summary, perhaps the most important conclusion from the present research is that the instruments currently available for assessing individual differences in mental imagery ability are not completely satisfactory. Definitive conclusions on the functional significance of these differences, in terms of their ability to predict performance on criterion tests of imagery ability, require the development of new measures whose items reflect a wider range of difficulty levels, and that make clearer distinctions between the dimensions of vividness and controllability.

References

Ashton, R., & White, K. (1974). Factor analysis of the Gordon test of visual imagery control. Perceptual and Motor Skills, 38, 945-946.

Betts, G.H. (1909). The distribution and functions of mental imagery. Teachers' College Columbia University Contributions to Education, 26, 1-99.

Block, N. (1981). Imagery. Cambridge: MIT Press.

Chambers, D. & Reisberg, D. (1985). Can mental images be ambiguous? Journal of Experimental Psychology: Human Perception and Performance, 11, 317-328.

Ernest, C. (1977). Imagery ability and cognition: A critical review. Journal of Mental Imagery, 2, 181-216.

Finke, R. (1980). Levels of equivalence in imagery and perception. Psychological Review, 87, 113-132.

Gordon, R. (1949). An investigation into some of the factors that favour the formation of stereotyped images. British Journal of Psychology, 39, 156-167.

Hilgard, E.R. (1981). Imagery and imagination in American psychology. Journal of Mental Imagery, 5, 5-66.

Holt, R.R. (1964). Imagery: The return of the ostracized. American Psychologist, 12, 254-264.

Katz, A. (1983). What does it mean to be a high imager? In J. Yuille (Ed.), Imagery, memory, and cognition. Hillsdale, N.J.: Erlbaum.

Kihlstrom, J.F. (1981). Puzzles of imagery. Journal of Mental Imagery, 5, 42-44.

Kihlstrom, J.F., & Harackiewicz, J.M. (1982). The earliest recollection: A new survey. Journal ot Personality, 50, 134-148.

Kosslyn, S.M. Image and mind. Cambridge: Harvard University Press.

Lorenz, C., & Neisser, U. (1985). Factors of imagery and event recall. Memory and Cognition, 13, 494-500.

Marks, D.F. (1973). Visual imagery in the recall of pictures. British Journal of Psychology, 64, 17-24.

Marks, D. (1983). Mental imagery and consciousness: A theoretical review. In A. Sheikh (Ed.), Imagery: Current theory, research, and application. New York: Wiley.

Peterson, M.A., Kihlstrom, J.F., Rose, P.M., Glisky, M.L., & Harvey, E. (1989). Mental images can be ambiguous: Parts, wholes, and strategies. Manuscript in preparation, University of Arizona.

Reisberg, D., & Heuer, F. (1989a). The consequences of vivic imagery: An empirical handle on the function of phenomenal states? Unpublished manuscript, Reed College.

Reisberg, D., & Heuer, F. (1989b). Vividness, vagueness, and the quantification of visualizing. Journal of Mental Imagery, in press.

Richardson, J. (1980). Mental imagery and human memory. New York: St. Martin's.

Sheehan, P.W. (1967). A shortened form of Betts' Questionnaire upon Mental Imagery. Journal of Clinical Psychology, 23, 386-389.

Sheehan, P.W. (1978). Mental imagery. In B.M. Foss (Ed.), Psychology survey, No. 1. Pp. 58-70. London: Allen & Unwin.

Sheehan, P.W., Ashton, R., & White, K. (1983). Assessment of mental imagery.

In A.A. Sheikh (Ed.), Imagery: Current theory, research, and application. New York: Wiley-Interscience.

Shepard, R.N., & Cooper, L.A. (1982). Mental images and their transformations. Cambridge: MIT Press.

Tower, R.B., & Singer, J.L. (1981). The measurement of imagery: How can it be clinically useful? In P.C. Kendall & S.D. Hollon (Eds), Assessment strategies for cognitive-behavioral interventions. Pp. 119-159. New York: Academic.

White, K., Ashton, R., & Brown, R.M.D. (1977). The measurement of imagery vividness: Normative data and their relationship to sex, age, and modality differences. British Journal of Psychology, 68, 203-211.

White, K., Ashton, R., & Law, H. (1974). Factor analyses of the shortened form of Betts' Questionnaire upon Mental Imagery. Australian Journal of Psychology, 26, 183-190.

White, K., Sheehan, P.W., & Ashton, R. (1977). Imagery assessment: A survey of self-report measures. Journal of Mental Imagery, 1, 145-170.

Yuille, J.C. (1983). Imagery, memory, and cognition: Essays in honor of Allan Paivio. Hillsdale, N.J.: Erlbaum.

Table 1

Means and Standard Deviations for QMI Sub-cales

Subscale	M	SD
Visual	11.60	4.40
Auditory	12.58	4.64
Tactile	12.04	4.60
Kinesthetic	12.10	4.60
Gustatory	12.70	5.13
Olfactory	14.44	5.49
Organic	10.83	4.27

N = 2036

Table 2

Correlations among QMI and TVIC subscales

		Betts Vis	Aud	Tac	Kin	Gus	Olf	Org
Betts	Visual	---						
QMI	Auditory	.36						
	Tactile	.43	.54					
	Kinesthetic	.39	.54	.58				
	Gustatory	.36	.47	.59	.52			
	Olfactory	.33	.43	.53	.45	.55		
	Organic	.38	.39	.49	.49	.47	.48	
Gordon TVIC	Total	.25	.21	.22	.25	.20	.18	.19

N - 2029

Table 3

Factor Loadings of Items on Gordon's
Test of Visual Imagery Control

TVIC Item	QMI/TVIC Factor 8	9	10	11	VVIQ/TVIC Factor 1	4
01. Car standing in road	.23	.02	.23	.64	.63	.06
02. Car in color	.15	-.02	.83	.19	.74	-.06
03. Car in different color	.14	.21	.81	.10	.64	.14
04. Car upside down	.16	.59	-.00	.42	.44	.59
05. Car upright again	.14	.26	.11	.71	.72	.27
06. Car moving on road	.70	.11	.10	.05	.79	.12
07. Car climbing hill	.84	.12	.02	.14	.77	.22
08. Car climbing over top	.82	.16	.03	.10	.70	.29
09. Car crashing through house	.13	.77	.06	.02	.21	.80
10. Car moving with couple inside	.45	.16	.27	.11	.59	.29
11. Car fall into stream	.15	.77	.04	.06	.25	.83
12. Car dismantled	.01	.68	.08	.04	.16	.74

Author Notes

This research was supported in part by Grant #MH-35856 from the National Institute of Mental Health to John F. Kihlstrom, and in part by Grant #BNS-8810997 from the National Science Foundation to Mary A. Peterson. We thank Kevin M. McConkey for his comments on earlier drafts.

Footnotes

[1]The TVIC shares this problem. The 12 items of the original 3-point version have endorsement rates varying from 55.8% to 95.9% ($M = 77.12$, $SD = 13.25$), with most of the low endorsement rates accounted for by the few unusual or bizarre images. It was in an attempt to increase the dispersion of scores on the TVIC that we introduced the 5-point scale used in the VVIQ sample.

[2]In the factor analyses performed for this research, TVIC items loaded on the first factor of each unrotated solution, along with QMI or VVIQ items. In the rotated solutions, however, TVIC items showed highest loadings on their own unique factors.

Canonical Descriptions, Reference Frames, and Viewer Intention in Figure-Ground Organization

Mary A. Peterson

University of Arizona

Paper presented at the first meeting of the American Psychological Society, Alexandria, Virginia, June 10 -12, 1989.

Canonical Descriptions, Reference frames, and Viewer Intentions

in Figure-Ground Organization

Many current models of shape recognition assume that a structural description of a stimulus shape is formed after the parts of a shape have been identified by some parsing process such as locating the minima of curvature from the inside of a figure. Recognition follows a massively parallel search in which the structural description is matched through multiple orientation mappings to multiple canonical memory representations. The shape's identity and orientation are arrived at simultaneously through an iterative process involving excitatory and inhibitory links between the canonical shape units and the orientation mapping units as well as between units of the same type. Other theories of shape recognition such as Biederman's use viewpoint independent components.

In this paper, I will discuss evidence obtained from the reversals of figure-ground stimuli that is relevant to these models. I will show first that -- at least for figure-ground stimuli -- the part parsing mechanism does not always favor one side of a contour as figure. At times it may parse the contour from both sides simultaneously. Second, I will show that different orientation mapping units are activated serially rather than in parallel. Finally, I will provide some evidence regarding the mechanisms through which top-down variables can influence the process of shape recognition.

The experiments I will describe today were done in collaboration with Erin Harvey, Brad Gibson, Pat Rose, and Holly Weidenbacher. We used figure-ground stimuli to address questions regarding shape recognition for a number of reasons. First, since the contour in a figure ground stimulus can be fit to the region on either side, such stimuli were ideal for probing whether the contour could be matched against more than one canonical representation simultaneously. Second, figure-ground stimuli possess the important property that only one region can be seen as figure at a time. Consider the classic Rubin vase-faces stimulus shown in Figure 1

Figure 1

When the contour is assigned to the white region, the faces are seen and the black region appears contourless, becoming the ground against which the faces are viewed. When figure and ground reverse so that the black region appears to be figure, the goblet is seen, but the faces are not. The white region is now contourless and appears to continue behind the goblet. Inasmuch

as shape recognition is tied to figure-ground organization, studying the variables that contribute to figure-ground reversal may provide answers regarding processes of shape recognition.

For example, figure 1 is reversible, but figure 2 is less reversible.

Figure 2

In figure 1, the shapes seen when both the center and the surround appear to be figure are <u>prototypical shapes</u>; that is, they are good exemplars of basic level representations. In figure 2, were the surround to be seen as figure, it would not be a <u>prototypical</u> shape. That figure 2 is less likely to reverse than figure 1 suggests that the prototypicality of the shapes on either side of the figure-ground contour contributes to reversibility. Within a network theory, networks representing prototypical shapes might be expected to settle into equilibrium faster than non-prototypical shapes. Thus, were two shapes competing for the contour through a speeded search, the more prototypical shape would be more likely to be seen.

We set out to examine this phenomenon experimentally while holding stimulus variables constant and controlling subjective variables. To hold stimulus variables constant, we used the stimuli shown in Figure 3, which were designed in collaboration with Julian Hochberg.

Figure 3

Upon first viewing, most observers see the center black region as figure in these two stimuli and we chose them to be so biased. We were most interested in the likelihood that viewers would see the white surround as figure, however. As shown, the white region does not represent a prototypical shape. When the figures are rotated by 180 degrees, however, the surrounds do represent prototypical shapes. The surrounds in figure 3a represent hook-nosed men's faces in profile and the surrounds in figure 3b are silhouettes of standing women. (Hereafter I will refer to the orientation in which the prototypical shape can readily be seen in the surround as the upright orientation.)

Notice that rotating the figures by 180 degrees does nothing to change the closure, area, or symmetry of the center region. Nor does it change the parts into which the contour would be parsed by any of the current methods. Thus, stimulus variables remain constant with inversion. Only the prototypicality of the shapes varies.

We controlled subjective variables by using a procedure called the opposed-set procedure that I devised in collaboration with Julian Hochberg. In this procedure, observers view ambiguous figures with opposed sets while maintaining fixation on a fixation point. On one trial, they are asked to try to hold one interpretation. On another trial, they are asked to try to hold the alternative interpretation. For example, with these figures, they were asked to try to hold the center region as figure on one trial and to try to hold the surround as figure on another trial. The trials last 30 seconds during which time observers report which alternative they are perceiving by keeping one of two keys depressed for as long as they see a single percept and switching keys to report reversals. Thus for each trial, we have a record of the number of reversals that occur and of the individual durations that each key was depressed.

The data I will discuss today have been summarized by dividing the summed durations that each alternative was perceived across the 30 second trial by the number of reversals into that alternative. I will discuss the mean duration that an alternative was perceived when it was the intended interpretation (I) -- that is when viewers tried to hold that interpretation and when it was the unintended interpretation (U -- that is, when viewers tried to hold the alternative interpretation. We measure the effectiveness of the viewers' intentions as the increased time that a region was seen as figure when it was the intended figure over the time that it was seen as figure when it was the unintended region. Intention is considered to have an influence on perceived organization when this difference is significantly greater than zero.

(Let me just mention at this point that (where applicable) the effects I will discuss here are replicated in analyses of the total durations. In addition, other experiments have shown that neither eye movements nor response bias are plausible mechanisms through which intention operates in this procedure. I will be glad to discuss those experiments later if anyone has questions regarding those issues.)

In the first experiment I will report today, we used the opposed-set procedure with figure-ground stimuli to examine whether the surround was indeed more likely to be seen as figure when it represented a prototypical shape, once stimulus variables were held constant and subjective variables were controlled. Further, we examined an hypothesis regarding the mechanism through which subjective variables might be operating. If intention influences the match between the structural description and the canonical repesentation it might be more effective in the upright viewing condition.

Experiment 1

In part 1 of Experiment 1, sixteen subjects viewed these figures in the inverted orientation. They were all unaware of the prototypical interpretation available for the surround in the upright orientation.

Table 1

Table 1 shows that on "hold center" trials, the mean duration of the center-as-figure percept was 15 seconds. On "hold surround" trials, the mean duration of the surround-as-figure percept was 5.7 seconds. Both of these mean durations were longer than the mean duration of the corresponding responses when they were the unintended responses. Accordingly, I - U, the effect of intention was significantly greater than zero for both the center and the surround. Notice that the intention to hold the center as figure was more successful than the intention to hold the surround as figure. Also note that the mean duration of the center-as-figure interpretation was longer than that of the surround-as-figure interpretation in the unintended responses as well.

We divided the subjects into two groups for the second half of the experiment. One group continued to view the inverted stimuli without being told about the inverted prototypical interpretation for the surround. They served as our practice control group. The other group viewed the upright stimuli and were told of the prototypical interpretation for the surround. We compared the performance of these two groups in order to examine whether the surround was seen as figure longer when it was a prototypical shape and whether intentions to hold the surround as figure were more effective when it was a prototypical shape.

Table 2

Table 2 shows the change in performance from part 1 to part 2. Negative numbers indicate shorter mean durations in part 2; positive numbers indicate longer mean durations. The performance of the practice control group remained approximately stable from part 1 to part 2. The mean durations of the intended responses decreased slightly and nonsignificantly for both the center and the surround, as did the mean duration of the unintended responses. Intention effects remained approximately stable.

There were large differences in the performance of the experimental group in Part 2, however. These subjects now reported seeing the center as figure on "hold center" trials for significantly shorter mean durations than they had in part 1. (The difference is -12.6 seconds) They reported seeing the

surround as figure for significantly longer on "hold surround" trials than they had in part 1 (The difference is +7.1 seconds). Even in the unintended responses, the mean duration of surround-as-figure increased and the mean duration of center-as-figure decreased, although these changes were not as large as the changes in the intended responses. The intention measure shows that observers' intentions to hold the center as figure were less effective in part 2 than they had been in part 1, and their intentions to hold the surround as figure were more effective.

The increased duration of surround-as-figure reports for the prototypical surround on "hold surround" trials shows that the likelihood of maintaining a region as figure is higher when that region represents a prototypical shape. The decreased duration of center-as-figure reports on the "hold center" trials in the experimental group shows that when the prototypicality of the surround increases, the likelihood of obtaining that region as figure by reversal from the center-as-figure percept is also higher.

We account for the increased likelihoods of both obtaining and maintaining the prototypical region as figure by proposing the existence of intermittent unperceived opportunities for reversal at which time a parallel search is initiated for the shapes on both sides of the contour. Whichever representation first reaches some criterion activation level is perceived. If the currently perceived shape is reconfirmed faster than the alternative, it is maintained and no reversal occurs. If the alternative shape is confirmed faster than the currently perceived shape, a figure-ground reversal occurs.

We assume that the match of a prototypical shape to its canonical representation is faster than the match of a non-prototypical shape. With this assumption, this model can account for the longer durations that the surround is perceived to be figure on "hold surround" trials when it is a prototypical shape because it will be more likely to be maintained when an opportunity for a reversal occurs. We can also account for the shorter mean durations that the center is perceived to be figure on "hold center" trials when the surround is a prototypical shape because the surround will be more likely to be obtained when an opportunity for a reversal occurs.

There is an alternative explanation for the results of Experiment 1, however. Remember that the observers in the experimental group of Experiment 1 both knew of the prototypical interpretations for the surrounds and viewed upright shapes. Therefore, we cannot be certain whether the effects obtained there were attributable to the viewers' knowledge or to the speed of a parallel search. For example, once observers were aware of

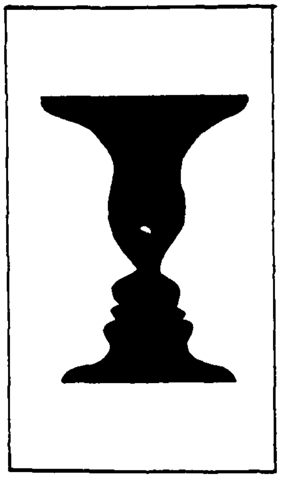


Fig. 1

Fig. 2

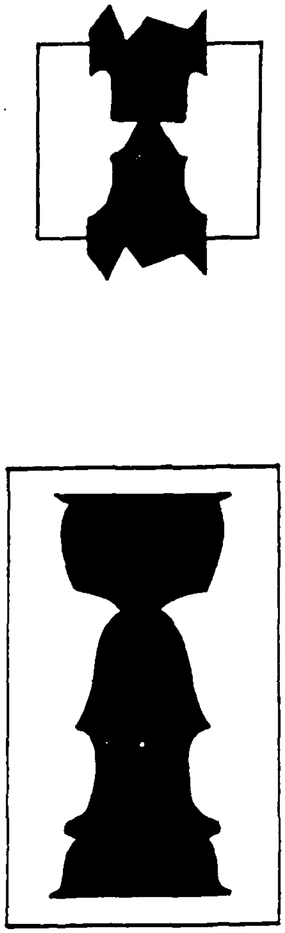


Fig. 3

MEAN DURATION OF A RESPONSE
PART 1
(SECONDS)

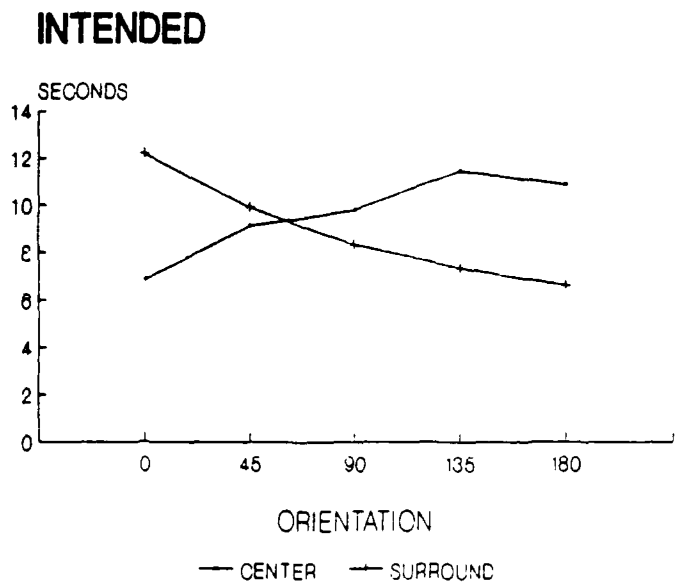
	CENTER	SURROUND
I	15.0	5.7
U	6.9	2.1
I-U	8.1	3.6

TABLE 1

Change in
MEAN DURATION OF A RESPONSE
PART 1—PART 2
(SEC)

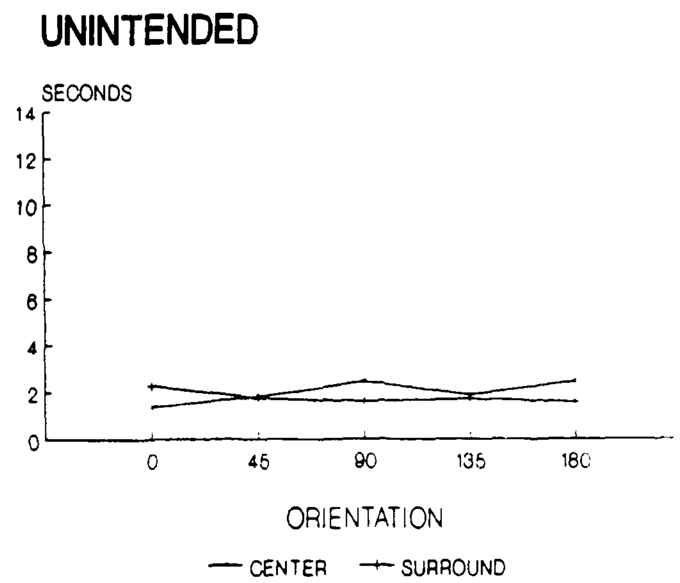
	PRAC (INV)		EXP (UPR)	
	C	S	C	S
I	-.4	-.3	-12.6	+7.1
U	-.8	-.2	-6.5	+1.7
I-U	+.4	-.1	-3.5	+5.4

TABLE 2



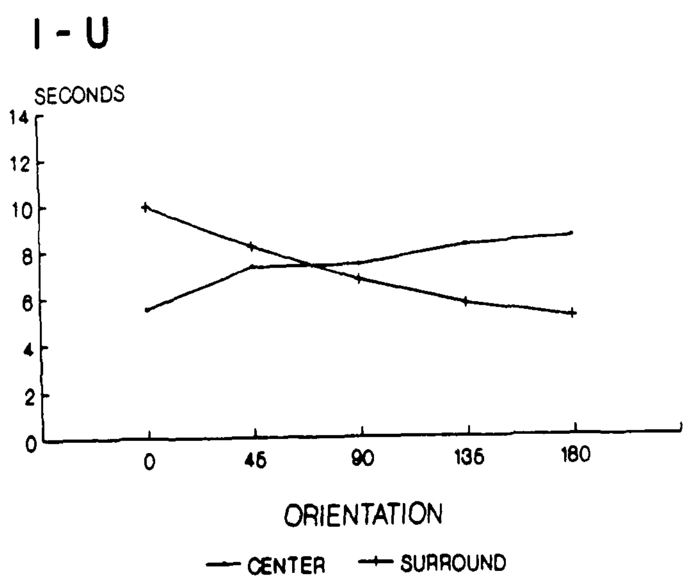
CENTER = CENTER SEEN AS FIGURE
SURROUND = SURROUND SEEN AS FIGURE

Fig. 4



CENTER = CENTER SEEN AS FIGURE
SURROUND = SURROUND SEEN AS FIGURE

Fig.5



CENTER = CENTER SEEN AS FIGURE
SURROUND = SURROUND SEEN AS FIGURE

Fig.6